

Influence of education, work, and conflicts on domestic abuse attitudes: A case study demonstrating data interoperability in Eastern Africa

Sula Sarkar (sanb0027@umn.edu)

IPUMS Centre for Data Integration, University of Minnesota, Minneapolis United States

Introduction and background:

Advances in big data and interdisciplinary research have made it imperative that one singular dataset or data source cannot meet the needs of holistic research. This compels the need for working with multiple datasets from varying sources and addressing issues related to data interoperability. The main objective of this study is to demonstrate using multiple datasets available from various sources to answer demographic questions. As examples, we use census microdata, data from the Demographic and Health Surveys (DHS), and data on violence and conflicts from the Armed Conflict Location and Event Data (ACLED). We investigate the effects of education attainment and labor force participation in the region as well as the effect of conflicts and riots on the domestic abuse attitudes of women in the area.

The United Nations 2030 Agenda for Sustainable Development Goals (SDG) proposes 17 goals and 169 targets that carry on the momentum generated by the Millennium Development Goals (UN, 2015ⁱ). The proposed framework for monitoring SDGs emphasizes the need for differing kinds of data and information from varying sources such as climate data, data on sustainable energy, pollution, population, and education data to name a few. Additionally, goals and target measurements need disaggregation at various demographic and social groups and at different levels of sub-national geography. This study demonstrates data interoperability at sub-national levels of geography. We answer two imperative research questions: first, whether educational attainment influences domestic abuse attitudes. For the purposes of this abstract, we only present results for the first research question. We use data from Kenya microdata census (2009) and Kenya Demographic and Health Survey (DHS) 2008. The second question we answer is whether educational attainment along with labor force participation rates, and violence and conflicts in the region influence domestic abuse attitudes. For the final paper, we will portray a multi country analysis from several countries in eastern Africa¹.

Data:

There are several datasets used in this study all interconnected to each other by a common thread - location or geography. For some datasets, geography is broad, like regions (equivalent to the primary level of geography in the respective country) in DHS data and census data. For some other datasets,

¹ Countries included in the final paper will include census and DHS data from Kenya (2009 census and 2008 DHS); Tanzania (2012 census and 2010 DHS); Uganda (2014 and 2011); Malawi (2008 and 2010); Zambia (2010 and 2013); Mozambique (2007, and 2011); Zimbabwe (2002 to 2010); Rwanda (2012 and 2010); Burundi (2008 and 2010). Data from ACLED will provide information on violence and riots from 2009 to 2013.

geographical information is specific, like location of conflicts from the ACLED project or the GPS clusters from DHS data.

Harmonized census data from IPUMS: Harmonized microdata census data are available from the Integrated Public Use Microdata Series-International (IPUMS-International). IPUMS disseminates high-precision census microdata samples from around the world including sample data for more than 550 censuses and surveys from 104 countries including census data from all the countries included in this analysis except Burundi. IPUMS makes a significant contribution to population research by optimizing data for cross-temporal and cross-national comparative analyses where variables are harmonized across IPUMS samples so that coding is always consistent across countries and samples. The data include variables representing a broad range of population characteristics, including fertility, life-course transitions, migration, disability, labor-force participation, occupational structure, education, ethnicity, and household composition (Ruggles et al. 2003; Sobek et al. 2011ⁱⁱ). Census data being high-density samples facilitates representation of a subset of population at lower administrative levels of geography. IPUMS provides **GIS boundary files** at the first and second administrative level of geography for the 104 countries in its collection. Additionally, the GIS files hold space constant, so that administrative boundaries are consistent through time making change over time studies meaningful.

Datasets from The DHS Program and IPUMS-DHSⁱⁱⁱ: DHS have been an invaluable source of information about the health and well-being of women, children, births, men and on all members of randomly selected households in low and middle-income countries over the past thirty years. This study uses domestic abuse attitudes and perception variables asked to a selected set of women of childbearing age (15 to 49). IPUMS-DHS integrates the DHS data to a common platform and makes it convenient for researchers to use. We include all women, of childbearing age, in our study who state that wife beating is justified if a woman 1) argues, 2) burns food, 3) goes out without telling the husband, 4) refuses to have sex, or 5) neglects children. This study also focuses on two kinds of spatial information from DHS. The first being the integrated DHS geographic regions available from IPUMS-DHS that hold space constant over timeⁱⁱⁱ. The second uses GPS clusters as locations of these women^{iv}. The GPS clusters² provide the coordinates for groups of households that participated in the DHS survey.

Data on riots, violence, protests are available from the ACLED project^v. The project records the precise geographical location, date, and type of event. An event in this case could be riots, violence against civilians, explosions, remote violence, battles, or protests. ACLED records events across Africa, South Asia, Southeast Asia, the Middle East, Central Asia and the Caucasus, and Southeastern and Eastern Europe and the Balkans. Dataset ranges from 1997 to present in parts of Africa and 2013 to present in parts of Asia. For the purposes of this study, the exact location and frequency of riots and violence are used as contextual information to study domestic violence attitude outcomes.

² GPS clusters are displaced 0-2 kilometers in urban and 0-5 kilometers in rural areas with 1% rural clusters displaced up to 10 kilometers to ensure respondent confidentiality. We account for displacement (as required) by creating spatial buffers around each GPS cluster.

Methods and preliminary results:

For the purposes of this abstract, we display a few preliminary results from Kenya, mainly answering our first research question - whether educational attainment influences domestic abuse attitudes. The focus of the final paper will be additional countries from Eastern Africa mentioned earlier (see footnote 1, page 1) and addressing both research questions. We use census microdata samples and calculate educational attainment at the first administrative level of geography. The sampling frame for DHS data in most countries is regions, which are usually the first-administrative level of geography. There are eight provinces in Kenya in both the datasets. With similar spatial footprints, numbers from both datasets can be used in conjunction with one another. Figure 1 illustrates stacked bar graphs with levels of education and domestic abuse for all the provinces of Kenya. Nairobi has higher educational attainment and a low percentage of women who perceive domestic abuse is justified. Western province is on the opposite end of the spectrum for domestic abuse perceptions, although it has a high enough primary education completion rate.

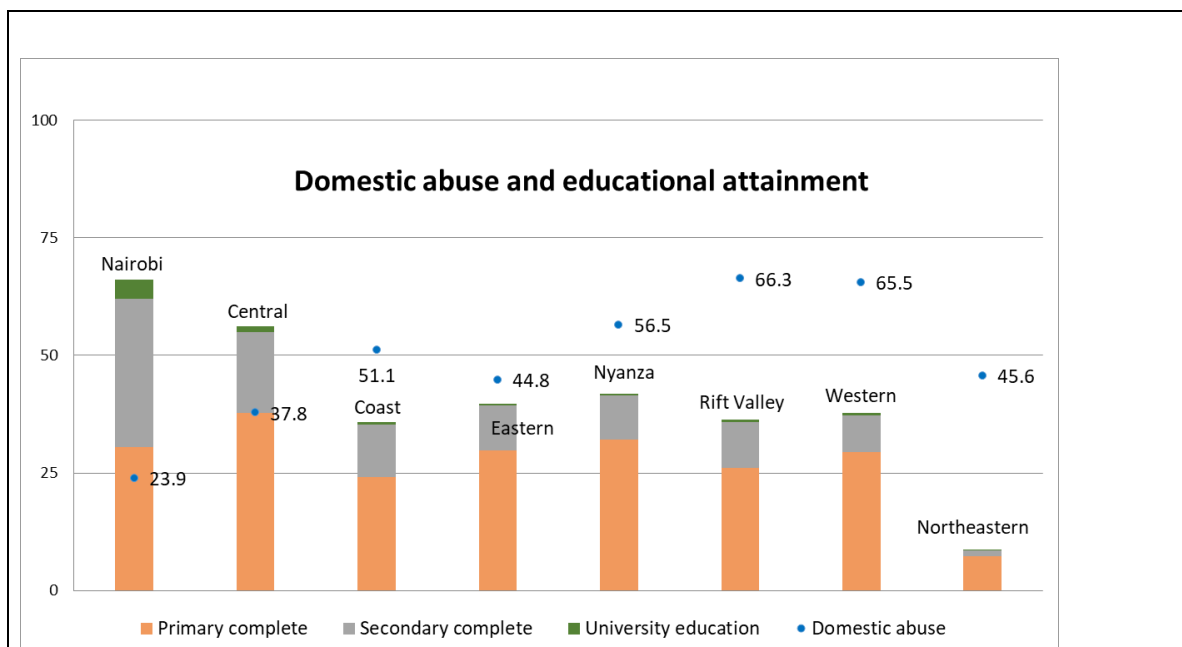


Figure 1: Comparing percentage women who think domestic abuse is justified to educational attainment (primary, secondary, and higher education) at the primary level of geography.

As we move into lower levels of geography (districts), we have enough cases to demonstrate educational attainment at the district level based on census samples. We take advantage of GPS cluster data from DHS, where we can locate the respondent. In Figure 2, we display a hot-spot analysis that identifies spatial clustering of women who perceive domestic abuse to be justified or not. A high Z score and small P value for a group of clusters indicates a significant hot spot (pink in Figure 2) whereas a low negative Z score and small P value indicates a significant cold spot (green in Figure 2). From Figure 2, although the cold spot is in an area (Nairobi in south central Kenya) where there is greater educational attainment,

there are parts of the country (such as Kisumu area in the Western province) which have higher educational attainment and a significant hot-spot. Higher primary, secondary, or university education completion rates are not necessary the areas where women's perception to wife beating is justified.

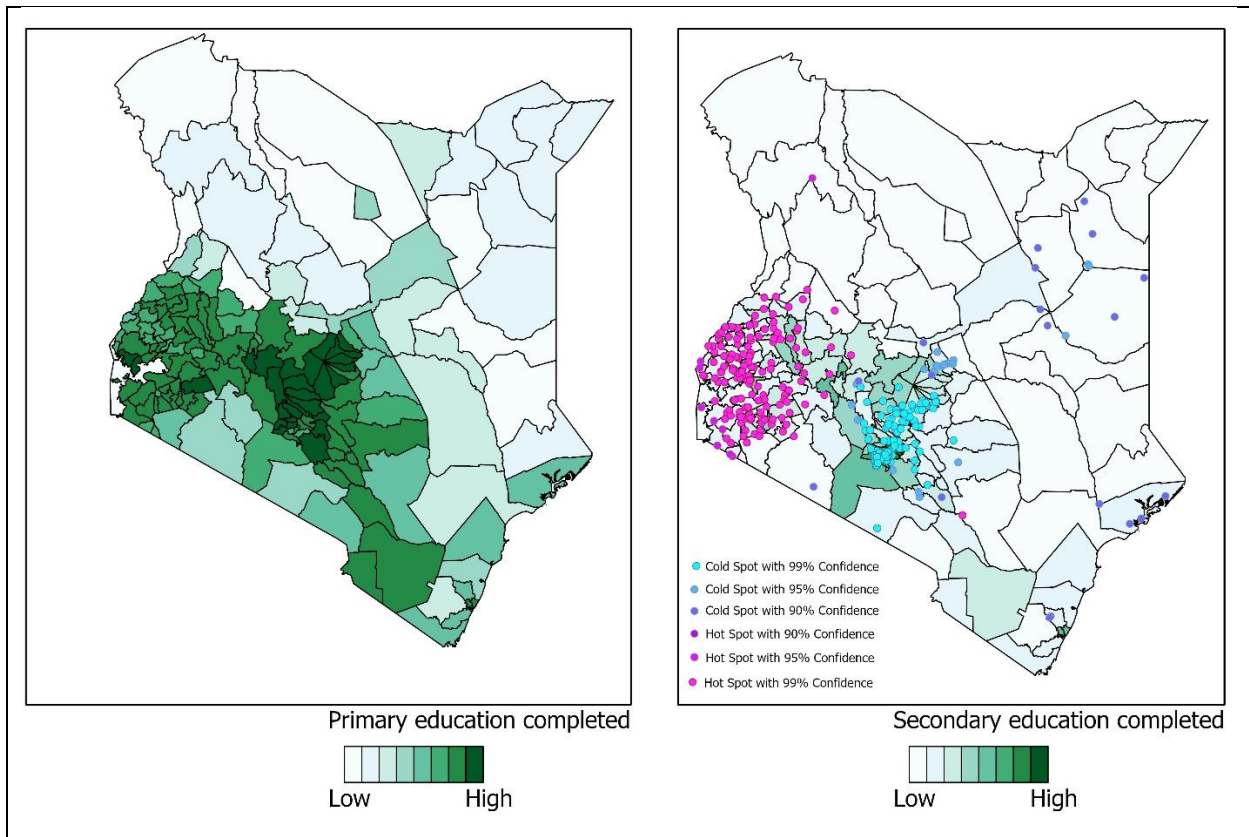


Figure 2: Comparing primary and secondary level of educational attainment to spatial clusters of women of childbearing age who perceive domestic abuse to be justified in Kenya. The color gradient used to visualize educational attainment is similar for both the maps above. A darker green color means higher primary or secondary education completion and a lighter shade of green signifies lower completion rates. The spatial clustering shows hot spots (in pink) meaning the women believe that domestic abuse is justified, and a spatial clustering of cold spots (blue) mean women do NOT perceive domestic abuse to be justified.

Summary and future work: This abstract portrays the use of multiple datasets available from differing sources and demonstrates how they can be used in conjunction with one another. Datasets are available at various geographical scales starting from the national level, to the first and then second administrative level of geography down to the approximate location of a person or event, like GPS clusters. Researchers need to be cautious about comparing results from similar spatial footprints across different datasets and the sampling frame for each dataset when working with lower resolution geographies. We visualize the association between education attainment at the primary and secondary level of geography to domestic abuse perceptions of women. We conclude that in Kenya, high education completion rates are not necessary the areas where women's perception to wife beating is justified. Further analysis is needed to

make a statistically significant conclusion. Our final paper will explore domestic violence perceptions in the eastern African region and take into consideration other factors like employment and look at the violence in the area.

References:

ⁱ United Nations General Assembly (2015). Transforming our world: The 2030 agenda for sustainable development. Retrieved from:
http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

ⁱⁱ Ruggles, Steven, Miriam L. King, Deborah Levison, Robert McCaa, and Matthew Sobek. 2003. "IPUMS-International." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 36 (2): 60–65. doi:10.1080/01615440309601215.

Sobek, Matthew, Lara Cleveland, Sarah Flood, Patricia Kelly Hall, Miriam L. King, Steven Ruggles, and Matthew Schroeder. 2011. "Big Data: Large-Scale Historical Infrastructure from the Minnesota Population Center." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 44 (2): 61–68. doi:10.1080/01615440.2011.564572.

ⁱⁱⁱ KNBS and ICF Macro: Kenya Demographic and Health Survey 2008–09. 2010, Calverton, Maryland: Kenya National Bureau of Statistics and ICF Macro National Council for Population and Development [Kenya] and Institute for Resource Development/Macro Systems Inc. Kenya Demographic and Health Survey 1989 [Dataset]. Data Extract from KEIR03.SAV, KEHH03.SAV, KEKR03.SAV, KEBR03.SAV, and KEPR03.SAV. IPUMS Demographic and Health Surveys (IPUMS DHS), version 7, IPUMS and ICF [Distributors]. Accessed from <http://idhsdata.org>

^{iv} Incorporating Geographic Information Into Demographic and Health Surveys: A Field Guide to GPS Data Collection (2013), measure DHS

^v Raleigh, Clionadh, Andrew Linke, Håvard Hegre and Joakim Karlsen. (2010). Introducing ACLED-Armed Conflict Location and Event Data." *Journal of Peace Research* 47(5) 651-660.