Education and region effects on Philippines teenage fertility in 1998*

Zhongyu Huang

April 6, 2022

Abstract

Teenage fertilities induce health risks and life impacts on both young mothers and their children. Philippines, one of the countries that have the highest adolescent birth rates, the teenage pregnancy toll is still rising here. This paper uses data collected by the National demographic and health survey in 1998 to investigate how background characteristics affect the teenage fertility rates in the Philippines. Evidence has shown that most teenage mothers received less education and lived in rural areas.

1 Introduction

According to the World Health Organization, an estimated 12 million girls aged 15–19 years give birth every year in developing countries(*Adolescent Pregnancy* 2020). There are at least 777,000 girls aged under 15 years becoming pregnant yearly(*Adolescent Pregnancy* 2020). Young individuals are expected to give birth after they successfully step into adulthood since early fertility leads to social and health-related impacts. It is not hard to imagine the difficulties that young mothers would face as they are economically dependent. Moreover, in current society, pregnant teenagers are disadvantaged. Research has shown that it is more likely for them to experience violence within partnerships(Kaye Wellings 1999). Other than that, early fertility contributes to health risks such as hypertension, low birth weight, and a high probability of spontaneous abortion(*Adolescent Pregnancy* 2020). Thus, it is necessary and important to find ways to reduce teenage fertility.

This paper uses data collected by the National Demographic and Health Survey to seek patterns in the background information of teenage mothers from the Philippines in 1998((NSO) 1999). Data related to fertility, child health, and family planning were reported in the DHS final report((NSO) 1999). We will use R(R Core Team 2020) to phase the Table 3.10 Teenage pregnancy and motherhood on page 45 of the pdf file to a useable data frame. By using appropriate graphs, we will analyze the relationships between teenage fertility and region, education level, and residence respectively.

Results indicate that the percentage of women who are mothers increased greatly from 0.3 for the 15-years-old girls to 17 for those who were 19. Differences in residence, region, and education level lead to quite different rates. In the Philippines, rural regions of Western Mindanao, Eastern Visayas, and Cagayan Valley had the highest proportions of early fertility. The level of education had a significant effect on this rate as well. Individuals who received more education understand more of the impact caused by early fertility. Thus, we could conclude that providing more access to education would be an efficient way to reduce teenage fertility.

The rest of the paper would be: Section "Data": introduce the survey and data. Section "Results": graphical presentation of results. Section "Discussions": discussions on what we find. Section "Weaknesses and next steps": talk about limitations involved in our analysis.

*Code and data are available at: https://github.com/zhongyuhuang/DHS_Analysis-.git

2 Data

Our data is collected by the Philippines National Demographic and Health Survey (DHS) in 1998. The DHS surveyed 13,983 women aged 15-49 nationwide by using questionnaires and interviews((NSO) 1999). It was designed to analyze potential trends in fertility, children's health, and family planning knowledge and use((NSO) 1999). We focus on the Teenage pregnancy and motherhood table on page 45 and phase it to a data frame from a pdf file.(Figure 1) The table summarized the percentage of teenage mothers with different background information from 2924 teenagers((NSO) 1999). Mainly with these four characteristics: Age, Residence, Region, and Education. More detailed information about the dataset can be find in the Appendix A.

I created five variables: BG_inf stores the background information of teenagers, PT_Mom stores the percentage who are mothers, PT_Prg stores the percentage who are pregnant with their first child, PT_bear stores the percentage who have begun child-bearing, and NumOfTeen stores the total number of teenagers in that category. Non-surprisingly, the rate of teenage pregnancy increased linearly with age(Figure 2). The percentages of teenagers who are mothers or pregnant in the different backgrounds are linked in some ways. We see that regions had high rates are more likely less urbanized and with less access to education.

Table 3.10 Teenage pregnancy and motherhood

Percentage of teenagers age 15-19 years who are mothers or pregnant with their first child, by selected background characteristics, Philippines 1998

Background characteristic	Percentage who are: Mothers	Pregnant with first child	Percentage who have begun child-bearing	Number of teenagers
Age				
15	0.3	0.2	0.5	624
16	0.8	0.8	1.6	613
17	3.6	1.4	5.0	589
18	7.8	2.6	10.5	602
19	17.0	4.2	21.3	497
Residence				
Urban	3.4	1.3	4.7	1,701
Rural	8.3	2.4	10.8	1,223
Region				
Metro Manila	2.8	0.9	3.7	605
Cordillera Admin.	2.3	2.3	4.6	46
Ilocos	3.3	1.3	4.6	147
Cagayan Valley	5.0	5.6	10.6	106
C. Luzon	5.1	1.0	6.2	289
S. Tagalog	7.2	2.7	9.9	360
Bicol	8.4	1.4	9.8	135
W. Visayas	4.6	1.0	5.6	233
C. Visayas	4.3	1.0	5.3	228
E. Visayas	8.5	3.5	12.0	102
W. Mindanao	8.3	4.4	12.6	112
N. Mindanao	7.0	1.2	8.2	106
S . Mindanao	6.4	2.3	8.7	206
C. Mindanao	5.7	1.3	7.0	98
ARMM	11.4	1.7	13.1	83
Caraga	7.8	1.3	9.2	68
Education				
No education	17.3	0.0	17.3	15
Elementary	11.5	4.4	15.9	471
High school	4.5	1.2	5.7	1,962
College or higher	3.0	1.5	4.5	476
Total	5.5	1.8	7.2	2,924

Figure 1: The original table in DHS final report

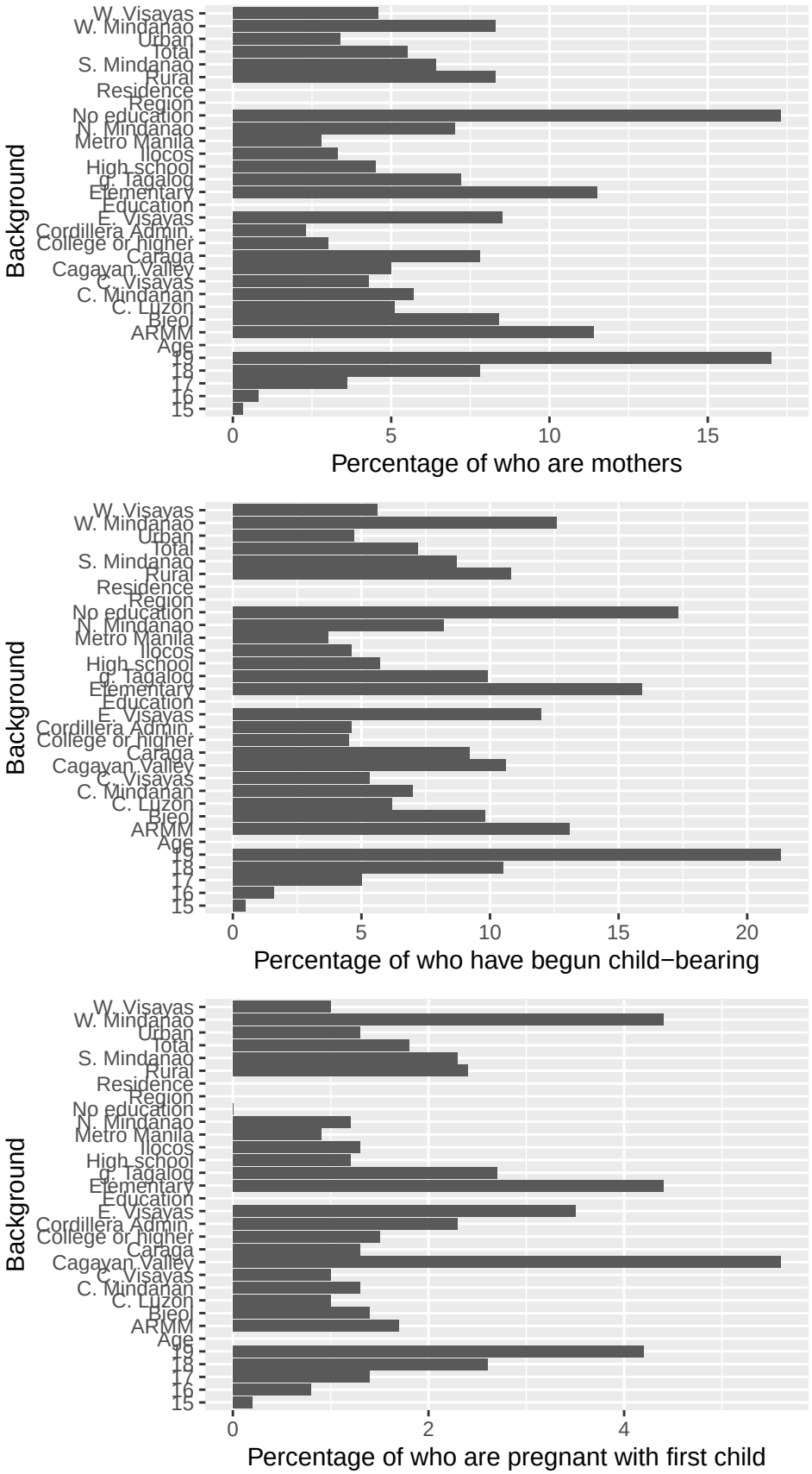

Figure 2: Visual presentation of variables

3 Results

Although teenage mothers only take up a small proportion of teenagers in the Philippines, we can see a significant difference between the backgrounds of these individuals. As for the residence type, proportions of rural teenage fertility are almost twice larger than those in urban residences(Figure 3). Most of them had already begun child-bearing(10.8%). Individuals aged from 15 to 19 are supposed to attend at least high school. However, we find out that most of them received no education(34.6%) or only finished elementary school(31.8%)(Figure 5). We also observed a large increase in the proportion who are mothers from 3% among teens who attended college or higher to 17% among those with no education. By comparing teenage fertility by regions, we can see western Mindanao, ARMM, and the eastern Visayas are regions that share the highest proportions of teenage mothers(Figure 4). Consistent with the rural-urban differences, these three regions are less urbanized compared to others. Non-surprisingly, Metro Manila, the national capital region who are urbanized and with more access to education, corresponds to the lowest teenage fertility.

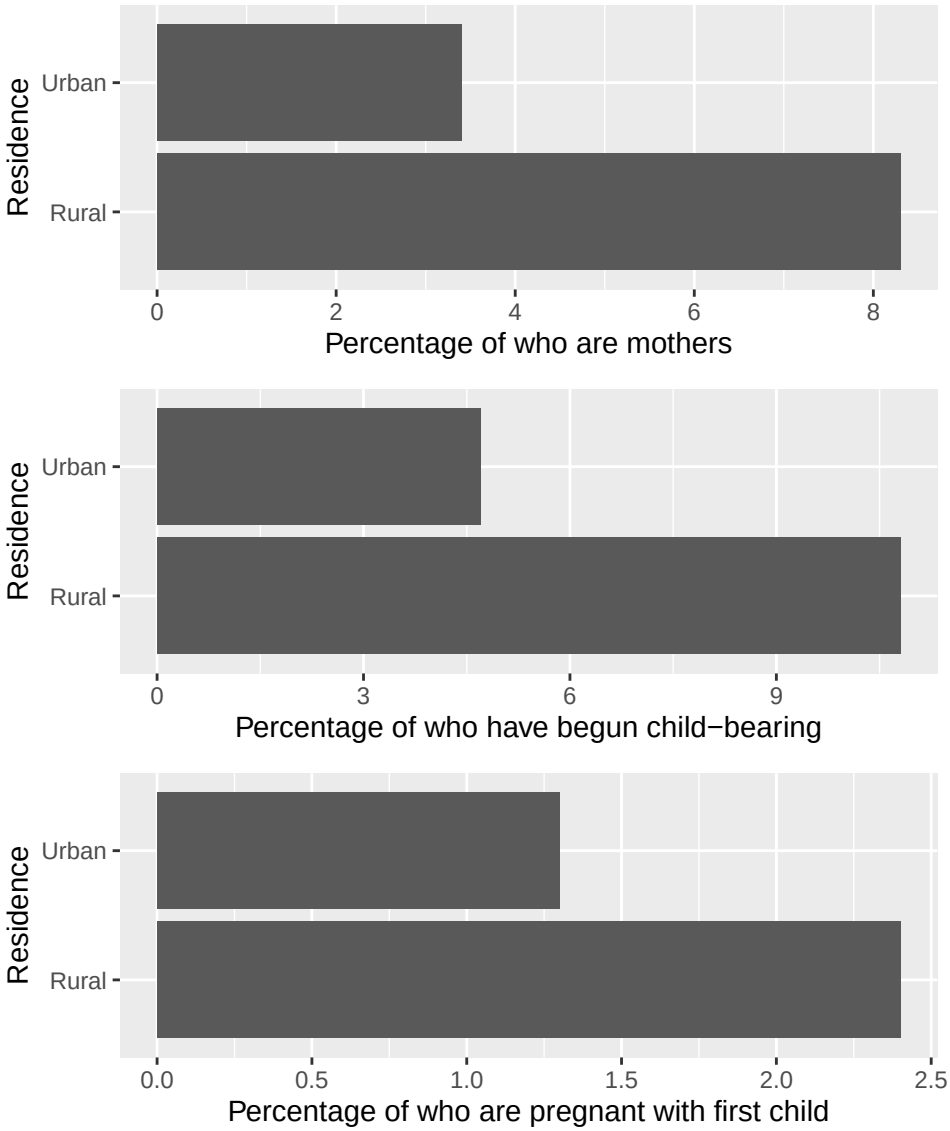

Figure 3: Percentage of teenage pregnancy and motherhood by residence

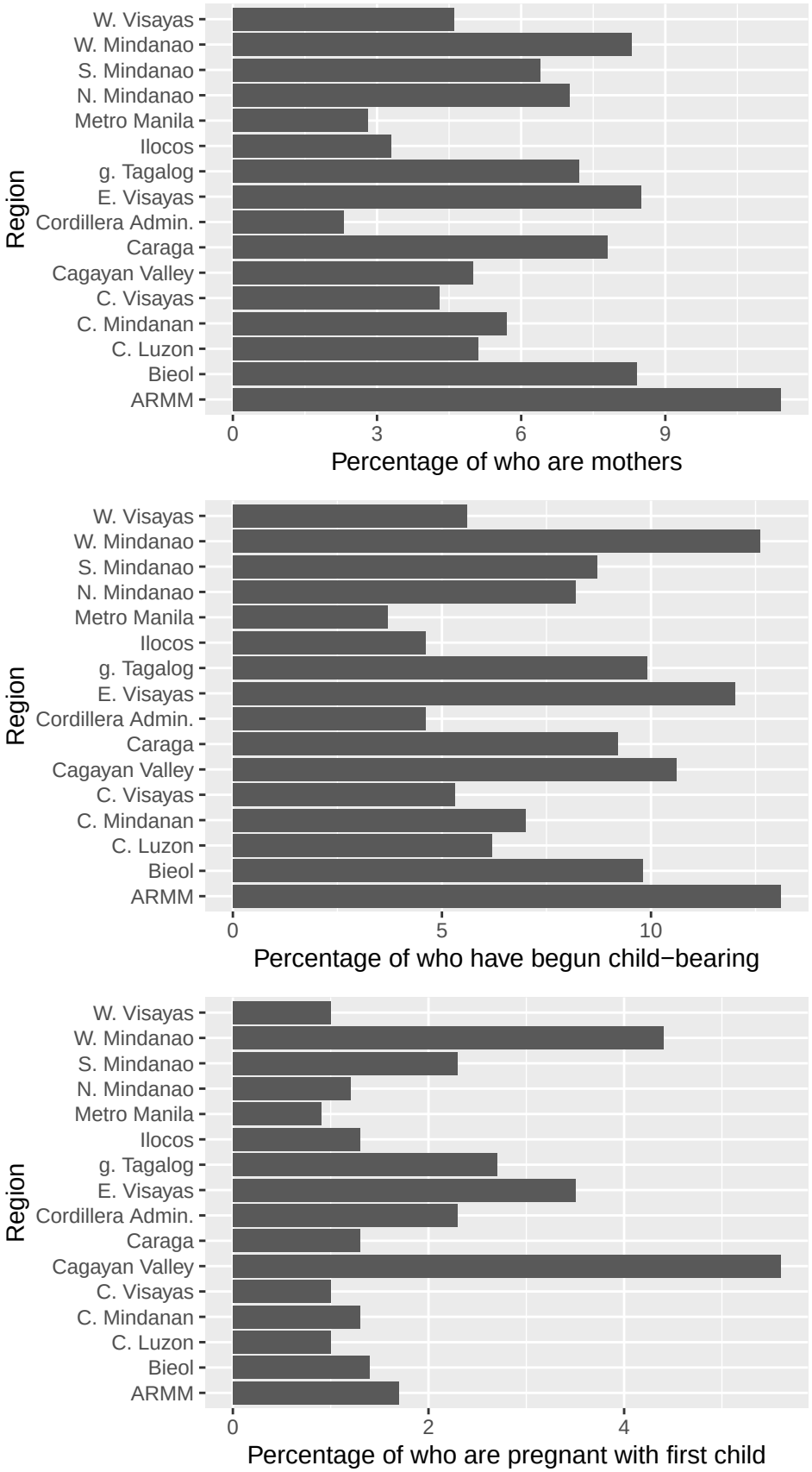

Figure 4: Percentage of teenage pregnancy and motherhood by region

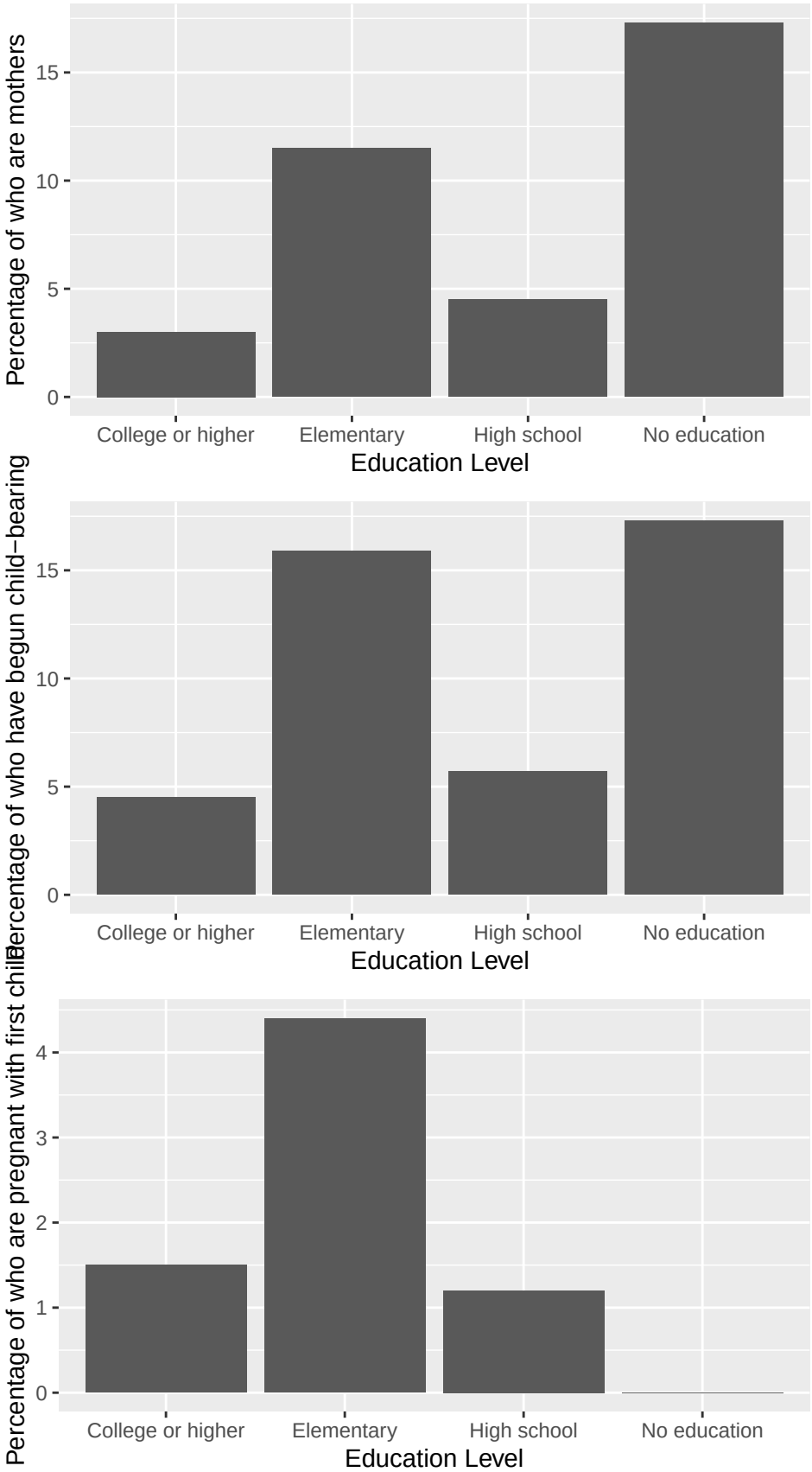

Figure 5: Percentage of teenage pregnancy and motherhood by education level

4 Discussion

4.1 Education matters

From the pattern discovered above, we can conclude with confidence that the effect of education on teenage fertility is significant. Individuals who have greater access to education could be more knowledgeable on the importance of safe sex. Education in the Philippines is a product of generations of colonialism(Joel M. Durban 2012). Thus, the current education system is still not impeccable and has issues remain. According to the research, there are only about 70% of pupils who enrolled in elementary school and finished the curriculum(Joel M. Durban 2012). This number even decreases to 30% for those enrolled in secondary or higher institutions(Joel M. Durban 2012). Well, where did these children go? Poverty is one of the major issues in the Philippines society. Due to the wide gap between the rich and the poor in the Philippines, children in poor families are forced to work to provide needs for their lives(Fernandez and Abocejo 2014). More concerning, children who attend school have not received a high-quality education. According to research, poor facilities are resulted from that the education system does not receive much budget from the Philippines government(Joel M. Durban 2012). Especially schools in the rural areas only received limited support from the government. Hence, teenage fertility is not a stand-alone issue in the Philippines. It is rather a complex product of multiple issues that exist in society.

4.2 Religion plays a role?

When investigating the difference in the proportion of teenage fertility by region, I discovered that the regions with the highest percentages are Eastern Visayas, the autonomous region, and Cagayan Valley. The corresponding religions for the majority of the households are Roman Catholics and Muslims. Would religion play a role in teenage fertility? According to the report conducted in 2004, there was not enough evidence to suggest that Muslim Pronatalism was linked to the lower power of Muslim women(S. Philip Morgan 2004). However, they found that compared to Muslim and non-Muslim communities in South Asia, the Muslim community had more children and tend to have more children(S. Philip Morgan 2004). As for Roman Catholics communities, around 82% of the population in the Philippines is Roman Catholic(Jones and Nortman 2004). Similar to Muslim communities, the research observed high birth rates of Roman Catholics in developing countries(Jones and Nortman 2004). We can not determine whether religion impacts teenage fertility without further investigation.

4.3 Conclusion and limitation

It is evident that teenage fertility could be influenced by residence, region, and education level in the Philippines 1998. These differences might reveal deeper factors related to this issue. This analysis is limited to women aged from 15 to 49 years. Moreover, we need more data from years to verify if these differences existed for years and not by chance. The hypothesis that religion might influence teenage fertility is weak without further studies of the histories of each religion. There are indeed more factors that should be taken into consideration to explain teenage fertility in the Philippines or any other country. Potential variables including income, access to abortion, and social custom might also linked to teenage fertility.

Appendix

A datasheet

Motivation

1. *For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.*
    - The dataset was created to analyze demographic data on fertility, family planning, and maternal and childhealth in Philippine.
2. *Who created the dataset (for example, which team, research group) and on behalf of which entity (for example, company, institution, organization)?*
    - The dataset was created by the National Statistics Office in collaboration with the Department of Health (DOH), the University of the Philippines Population Institute, and other concerned agencies in the Philippine government.
3. *Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.*
    - The creation of the dataset is funded by the U.S. Agency for Ituemational Development and the DOH.
4. *Any other comments?*
    - This dataset is the product of the 1998 NDHS which is part of the worldwide Demographic and Health Surveys (DHS) program.

Composition

1. *What do the instances that comprise the dataset represent (for example, documents, photos, people, countries)? Are there multiple types of instances (for example, movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.*
    - The 1998 NDHS sample is a sub-sample of the new master sample of the Integrated Survey of Households (ISH) of the NSO.
2. *How many instances are there in total (of each type, if appropriate)?*
    - There are 13,983 instances in total.
3. *Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (for example, geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (for example, to cover a more diverse range of instances, because instances were withheld or unavailable).*
    - The dataset collects information from all sixteen regions of the country. It can be considered as containing all possible instances.
4. *What data does each instance consist of? "Raw" data (for example, unprocessed text or images) or features? In either case, please provide a description.*
    - Each instance consists of raw data(categorical and numerical information).
5. *Is there a label or target associated with each instance? If so, please provide a description.*
    - The dataset targets women aged 15-49.

6. *Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (for example, because it was unavailable). This does not include intentionally removed information, but might include, for example, redacted text.*
    - There were about 2.8 non-response of the survey.
7. *Are relationships between individual instances made explicit (for example, users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.*
    - The relationships were not explicit.
8. *Are there recommended data splits (for example, training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.*
    - There were no recommended data splits.
9. *Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.*
    - There were no errors, sources of noise, or redundancies in the dataset.
10. *Is the dataset self-contained, or does it link to or otherwise rely on external resources (for example, websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (that is, including the external resources as they existed at the time the dataset was created); c) are there any restrictions (for example, licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.*
    - The dataset is self-contained.
11. *Does the dataset contain data that might be considered confidential (for example, data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? If so, please provide a description.*
    - The dataset does not contain confidential data.
12. *Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.*
    - The dataset does not contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety.
13. *Does the dataset identify any sub-populations (for example, by age, gender)? If so, please describe how these subpopulations are identified and provide a description of their respective distributions within the dataset.*
    - The datasets identify woman aged 15-49.
14. *Is it possible to identify individuals (that is, one or more natural persons), either directly or indirectly (that is, in combination with other data) from the dataset? If so, please describe how.*
    - It is not possible to identify individuals either directly or indirectly.
15. *Does the dataset contain data that might be considered sensitive in any way (for example, data that reveals race or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)? If so, please provide a description.*
    - The dataset contain no such data.
16. *Any other comments?*
    - No.

Collection process

1. *How was the data associated with each instance acquired? Was the data directly observable (for example, raw text, movie ratings), reported by subjects (for example, survey responses), or indirectly inferred/derived from other data (for example, part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.*
    - The data associated with each instance was reported by subjects from responding to survey and interviews. The organization did not mention if they verified the data.
2. *What mechanisms or procedures were used to collect the data (for example, hardware apparatuses or sensors, manual human curation, software programs, software APIs)? How were these mechanisms or procedures validated?*
    - The data was collect using manual human curation, including survey and interview.
3. *If the dataset is a sample from a larger set, what was the sampling strategy (for example, deterministic, probabilistic with specific sampling probabilities)?*
    - The dataset was a sub-sample of the new master sample of the Integrated Survey of Households (ISH) of the NSO.
4. *Who was involved in the data collection process (for example, students, crowdworkers, contractors) and how were they compensated (for example, how much were crowdworkers paid)?*
    - The process involved 35 trainers from NSO, DOH, UPPI, and POPCOM, 261 interviewers, 44 supervisors, and 43 field editors. I
5. *Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (for example, recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created.*
    - The data was collected in October 1997. It matched the creation timeframe of the data associated with the instances.
6. *Were any ethical review processes conducted (for example, by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.*
    - There was no ethical review processes conducted.
7. *Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (for example, websites)?*
    - I collected the data from https://dhsprogram.com/publications/publication-fr103-dhs-final-reports.cfm
8. *Were the individuals in question notified about the data collection? If so, please describe (or show with screenshots or other information) how notice was provided, and provide a link or other access point to, or otherwise reproduce, the exact language of the notification itself.*
    - Individuals were notified about the data collection.
9. *Did the individuals in question consent to the collection and use of their data? If so, please describe (or show with screenshots or other information) how consent was requested and provided, and provide a link or other access point to, or otherwise reproduce, the exact language to which the individuals consented.*
    - Individuals did consent to the collection and use of their data.
10. *If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses? If so, please provide a description, as well as a link or other access point to the mechanism (if appropriate).*

- There was no such mechanism provided.

11. *Has an analysis of the potential impact of the dataset and its use on data subjects (for example, a data protection impact analysis) been conducted? If so, please provide a description of this analysis, including the outcomes, as well as a link or other access point to any supporting documentation.*
    - There were no analysis of potential impact of the dataset and its use on data subjects.
12. *Any other comments?*
    - No.

Preprocessing/cleaning/labeling

1. *Was any preprocessing/cleaning/labeling of the data done (for example, discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remaining questions in this section.*
    - The data was extracted from the final report of 1998 DHS for Philippines.
2. *Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (for example, to support unanticipated future uses)? If so, please provide a link or other access point to the "raw" data.*
    - The raw data was saved in outputs/data/raw_data.csv
3. *Is the software that was used to preprocess/clean/label the data available? If so, please provide a link or other access point.*
    - R is used to preprocess/clean/label the data
4. *Any other comments?*
    - No

Uses

1. *Has the dataset been used for any tasks already? If so, please provide a description.*
    - The dataset has not been used for any tasks.
2. *Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point.*
    - https://github.com/zhongyuhuang/DHS_Analysis-.git
3. *What (other) tasks could the dataset be used for?*
    - The dataset could be used to analysis fertility preferences Philippines in 1998.
4. *Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? For example, is there anything that a dataset consumer might need to know to avoid uses that could result in unfair treatment of individuals or groups (for example, stereotyping, quality of service issues) or other risks or harms (for example, legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?*
    - There is nothing about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses.
5. *Are there tasks for which the dataset should not be used? If so, please provide a description.*
    - The dataset should only be used to analysis fertility preferences Philippines in 1998.

6. *Any other comments?*
    - No.

Distribution

1. *Will the dataset be distributed to third parties outside of the entity (for example, company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.*
    - The dataset wil not be distributed to third parties outside of the entity.
2. *How will the dataset be distributed (for example, tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?*
    - The dataset be distributed on github.
3. *When will the dataset be distributed?*
    - The dataset will be distributed on April 1, 2022.
4. *Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/ or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.*
    - No.
5. *Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.*
    - No third parties have imposed IP-based or other restrictions on the data associated with the instances.
6. *Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.*
    - There are no export controls or regulatory restrictions.
7. *Any other comments?*
    - No.

Maintenance

1. *Who will be supporting/hosting/maintaining the dataset?*
    - Zhongyu Huang
2. *How can the owner/curator/manager of the dataset be contacted (for example, email address)?*
    - Zhongyu.huang@mail.utoronto.ca
    - Additional information about the DHS program may be obtained by writing to: Macro International Inc., 11785 Beltsville Drive, Calverton, MD 20705-3119, USA (Telephone 301-572-0200, Fax 301-572-0999).
3. *Is there an erratum? If so, please provide a link or other access point.*
    - There is no erratum.
4. *Will the dataset be updated (for example, to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (for example, mailing list, GitHub)?*

- The dataset will not be updated.

5. *If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (for example, were the individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced.*

   - There were no such limitations.

6. *Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers.*

   - The older versions of the dataset will be supported, but not be maintained.

7. *If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.*

   - There is no such mechanism.

References

Adolescent Pregnancy. 2020. World Health Organiztion. https://www.who.int/news-room/fact-sheets/detail/adolescent-pregnancy.

Fernandez, Russell Cristian E., and Ferdinand T. Abocejo. 2014. "Child Labor, Poverty and School Attendance:evidences from the Philippines by Region." *CNU Journal of Higher Education* 8: 114–27. https://www.researchgate.net/profile/Ferdinand-Abocejo/publication/319505916_Child_Labor_Poverty_and_School_Attendance_Evidences_from_the_Philippines_by_Region/links/59aff691458515150e4ce656/Child-Labor-Poverty-and-School-Attendance-Evidences-from-the-Philippines-by-Region.pdf.

Iannone, Richard, and Mauricio Vargas. 2022. *Pointblank: Data Validation and Organization of Metadata for Local and Remote Tables*. https://CRAN.R-project.org/package=pointblank.

Joel M. Durban, Ruby Durban Catalan. 2012. "ISSUES AND CONCERNS OF PHILIPPINE EDUCATION THROUGH THE YEARS." *ASIAN JOURNAL OF SOCIAL SCIENCES & HUMANITIES* 1: 61–69. https://d1wqtxts1xzle7.cloudfront.net/27912681/ajssh2012(1.2-08)-with-cover-page-v2.pdf.

Jones, Gavin, and Dorothy Nortman. 2004. "ROMAN CATHOLIC FERTILITY AND FAMILY PLANNING: A Comparative Review of the Research Literature." *Studies in Family Planning* 1: 1–27. https://doi.org/10.2307/1964908.

Kaye Wellings, Anne Johnson, Jane Wadsworth. 1999. "Teenage Fertility and Life Chances." *Society for Reproduction and Fertility* 4: 184–90. https://doi.org/10.1530/revreprod/4.3.184.

(NSO), National Statistics Office. 1999. *National Demographic and Health Survey 1998*. Department of Health (DOH) [Philippines]; Macro International Inc. (MI).

Ooms, Jeroen. 2022. *Pdftools: Text Extraction, Rendering and Converting of PDF Documents*. https://CRAN.R-project.org/package=pdftools.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

S. Philip Morgan, Herbert L. Smith, Sharon Stash. 2004. "Muslim and Non-Muslim Differences in Female Autonomy and Fertility: Evidence from Four Asian Countries." *Population and Development Review* 28: 515–37. https://doi.org/10.1111/j.1728-4457.2002.00515.x.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Xie, Yihui. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. https://yihui.org/knitr/.

———. 2022. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.