Liberian Women Modern Contraceptive Use in 1986 and Its Influencing Social Factors*

Yiying Chen Jennifer Liu Zhiyi Liu

10 April 2022

Abstract

This paper mainly investigates how various social backgrounds of Liberian women affect their use of contraceptive methods. By data analysis, we conclude that age, level of education, number of current living children and tribes largely influence people' use of birth control. As a rising social issue, helping more people to realize the importance of modern contraception was a necessary step toward population control. From the result of this paper, we could identify groups of individuals that need more education on modern contraception and make relevant public policies to tackle this problem.

Contents

*Code and data avaliable at https://github.com/YiyingChen0523/Contraceptive_method.git

Introduction

The use of birth control plays a main role in women's decision of whether to have a child or not. The topic is essential for women's body autonomy with the rise of feminism worldwide. The method of contraceptives directly affects the effectiveness of avoiding pregnancy. Countries under development often face rapid changes in beliefs and practices. Liberia, one of the oldest republics in Africa, was studied to see the use of contraceptives among women and the percentage of methods considered modern. The study also collects data on respondents' information on social factors.

The Liberia Demographic and Health Survey (LDHS) was conducted on a national-level in 1986. The purpose of the survey was to give policymakers information regarding family planning and maternal health in the country of Liberia. The research was conducted in the format of survey, collecting data from a sample of 5,239 women aged from 15 to 49. Respondents are required to answer questions related to fertility, family planning, and maternal and child health (Statistics and Planning 1986).

The paper examines the use of contraceptives among each subgroup of age, and the proportion of modern contraceptive methods among all methods. We see that each group had a different percentage use of birth control and modern methods of birth control. We found that age factors contributed to a bimodal distribution of percentage of that group who is using contraceptives. The use was lowest for the age group 15-19. The percentage peaked with groups 25-29 and 45-49. For the number of children a woman has, more children is correlated with higher use of birth controls. A notable finding is that among all women who use contraceptives and do not have any children use modern methods, which might imply the effectiveness of modern contraceptives. Education level is also a social factor that exhibited increasing use of birth control as education level rises. Lastly, tribes in Liberia that have low use of contraceptives are Bassa, Gio and Gola. High-use tribes include Grebe, Kru/Sapo and Lorma. Among all the tribes, both Gola and Mandigo do not use any traditional methods.

The paper begins with the data section, including information on the source of data, sampling method, data collection and data characteristics. The results section displays the exact percentages among social groups. The paper divides the factors into age, number of living children, education level and the tribes. The discussion section provides deeper explanation and potential reasoning behind the results, along with limitations of the paper.

Data

This report is analyzed using R (R Core Team 2020), using tidyverse (Wickham et al. 2019) and dplyr (Wickham et al. 2020) packages. All the tables and graphs are created using ggplot2 (Wickham 2016) and the file is knitted using knitr (Franbois, Henry, and Miller 2021). The package janitor (Firke 2021), pdftools(Ooms 2022), purrr(Henry and Wickham 2020) are also used.

Data Source

This data is taken from the Liberia Demographic and Health Survey (LDHS) in 1986 and was implemented by the Ministry of Planning and Economic Affairs and the Institute for Resource Development. It was part of the worldwide Demographic and Health Surveys Programme (DHS) that aims to collect data on fertility, family planning, and maternal and child health. It has collected and analyzed over 400 surveys in more than 90 countries (Statistics and Planning 1986).

TABLE 4.6 PERCENTAGE CURRENTLY USING ANY CONTRACEPTIVE METHOD AND CURRENTLY USING ANY MODERN CONTRACEPTIVE METHOD AMONG CURRENTLY MARRIED WOMEN BY BACKGROUND CHARACTERISTICS, LIBERIA, 1986

Characteristic	Any Method	Any Modern Method	Characteristic	Any Method	Any Modern Method
Age			Education		
15-19	2.1	2.0	No schooling	2.8	2.5
20-24	5.4	4.9	Primary	7.3	6.6
25-29	7.7	6.7	Secondary or more	26.8	22.1
30-34	8.1	6.3			
35-39	5.2	5.1	Religion		
40-44	8.3	5.7	Christian	9.6	8.2
45-49	8.0	7.1	Muslim	3.5	2.8
			Traditional/Other	3.7	3.1
Living Children			None	3.5	3.3
None	2.5	2.5			
1	3.8	3.6	Tribe		
2	5.7	4.8	Bassa	4.0	3.3
3	6.4	5.8	Gio	3.6	3.1
4	9.7	7.3	Gola	2.8	2.8
5	6.5	5.5	Grebo	12.2	11.4
6 or more	12.1	10.6	Kpelle	3.7	3.5
			Krahn	4.0	4.0
Urban-Rural			Kru/Sapo	11.3	10.7
Urban	11.6	9.7	Lorma	12.1	9.8
Rural	3.4	3.1	Mandingo	1.5	1.5
			Mano	6.3	5.4
Region			Other/None	9.6	7.2
Sinoe	4.4	3.9			
Grand Gedeh	3.0	2.9			
Montserrado	12.0	9.7			
Rest of country	4.9	4.4	Total	6.4	5.5

The 1986 LDHS report is found to be in pdf format (Table 1) and in order to analyze it using R, page 44 has been converted into a table for analysis purposes.

Sampling

The survey has been oversampled as it also uses data collected in Sinoe and the Grand Gedeh Counties. Over 4500 enumeration areas were created based on the 1984 Population Census and 156 enumeration areas are covered in the census. They were selected with a probability proportional to the geographical size. EAs that were more than 3-4 hours away from a road were removed as they represent less than 3 percent of the total number of households. All households are then listed from the 156 enumeration areas and then sampled.

6006 households were selected for an interview, but only 5609 households were found due to factors such as the house has been emptied or destroyed. Out of the 5206 households that were found were interviewed, with a response rate of 90%. Out of these households, 5240 women were identified to be eligible and 98% were interviewed. They need to be between 15-49 years old and stayed in the households the night before.

Collection

LDHS uses two questions, one to list all the members of selected households and the other one to collect information needed. The field staff are all female and consist of 24 interviewers, where training was given. Data was collected from late February to July 1986 and then manually computed onto microcomputers at the Bureau of Statistics Office. Errors found in the microdata are resolved by referring to the questionnaire or by logic using other information given. Data quality remained uncertain as there is information that is undisclosed; only 42% gave the birth date.

Data Characteristics

The data distribution of the sample women is representative to the country's data collected in the 1974 census with two minor differences. The percentage of LDHS respondents in the 25-29 and 45-49 groups is higher than the census' proportion (17.7 % in 1974 census compared to 20.6% for 25-29). Moreover, there were only 28.5% of women living in urban areas in the 1974 census, compared to 43.2% for the survey.

Modern method of contraceptives refers to the use of pill, IDD, injection, vaginal methods, condom, female and male sterilization. Pill is the most used modern method among women and male sterilization is the less commonly used method.

Result

In this session, we divide the original table into sub-session to do a deeper analysis and conclude what are the factors that could largely affect people' use of contraceptive methods. The percentage currently using contraceptive methods varies a lot under different contexts. People who obtain specific features will obviously prefer using contraceptive methods or not. Overall, we select four features that could mostly represent the background information of an individual to analyze their choices of using contraceptive methods (Statistics and Planning 1986).

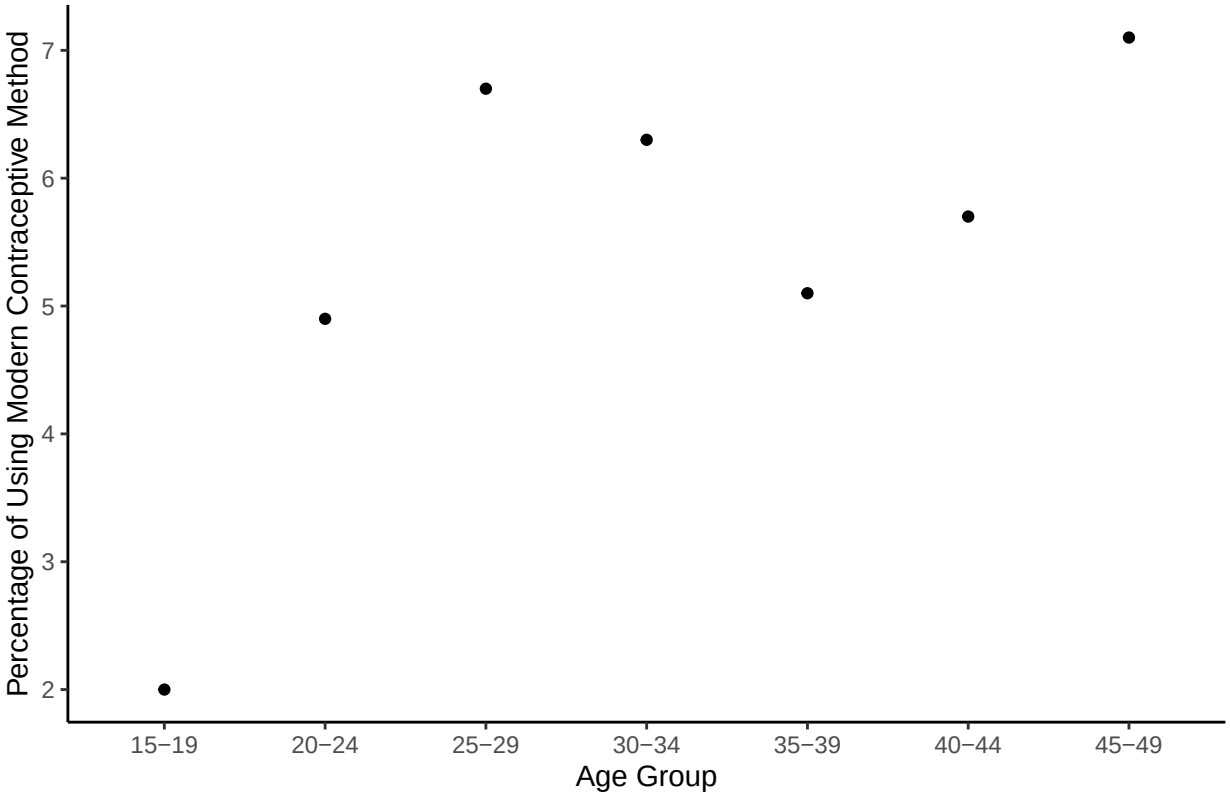
Age

The first factor was age, which was the one that could largely represent the fertility of individuals. From the trend of the data, we could see that people between the ages of 45-49 have the largest portion of people using modern contraceptive methods. They tend to have a stable living condition and control accidental pregnancies. The second age group that mostly used modern contraceptive methods was people between 25-29. Compared to younger groups, they have a better awareness of birth control and they know the scientific method. Conversely, for people between 15-19 years old, only 2% of them use modern contraceptive methods (Table 1).

Table 1: Summary By Age

Age	Any_Method	Any_Mordern_Method
15-19	2.1	2.0
20-24	5.4	4.9
25-29	7.7	6.7
30-34	8.1	6.3
35-39	5.2	5.1
40-44	8.3	5.7
45-49	8.0	7.1

Figure1: Percentage of Using Modern Contraceptive Method by Age



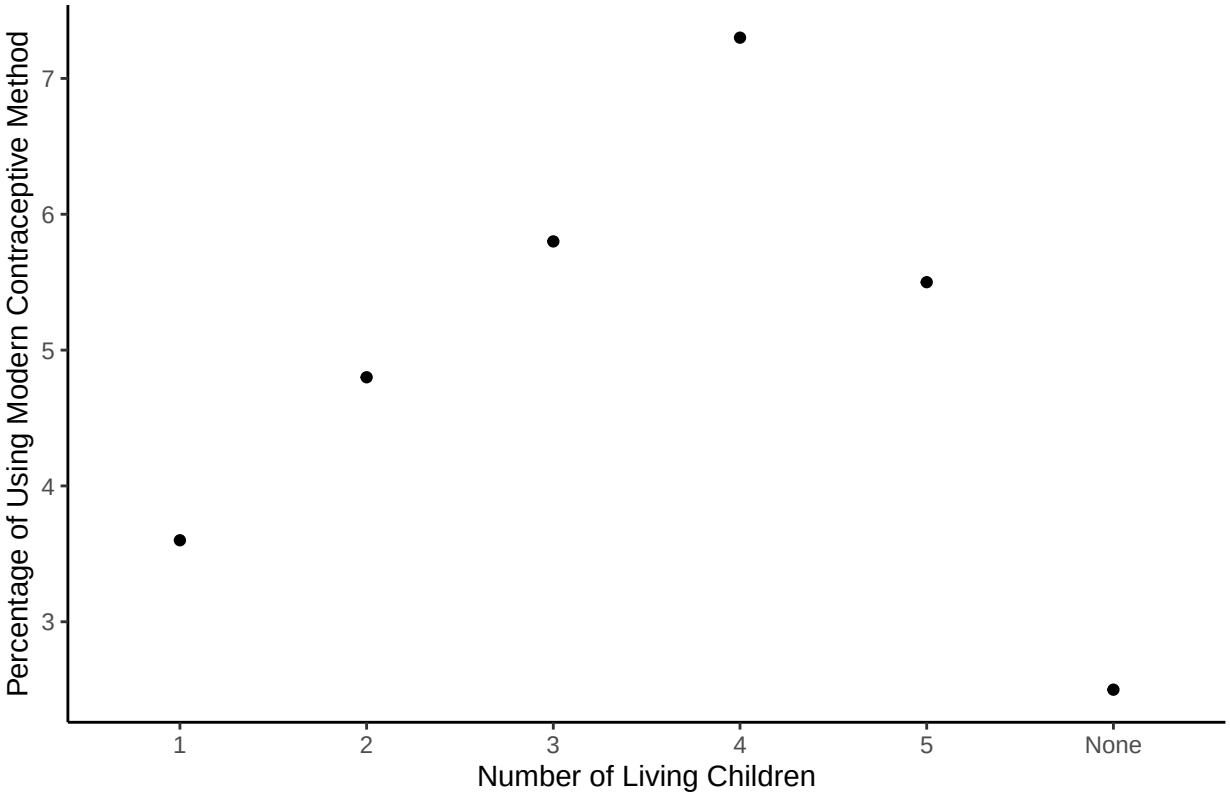
Living Children

Then, the current number of children was also a key factor for people to decide whether to use contraceptive methods. Most families prefer using modern contraceptive methods mostly. The distribution is left-skewed, indicating that people tend to use more contraceptives once they have more children. The peak is at when people have 4 children, 9.7% of the population in that group uses contraceptives and 7.3% uses modern methods (Table 2). The distribution of people who use modern methods also follows the same distribution of all methods. It makes sense because it is included in the any methods category (Figure 2).

Table 2: Summary By Living Children

Living_Children	Any_Method	Any_Mordern_Method
None	2.5	2.5
1	3.8	3.6
2	5.7	4.8
3	6.4	5.8
4	9.7	7.3
5	6.5	5.5

Figure 2: Percentage of Modern Method by Number of Children

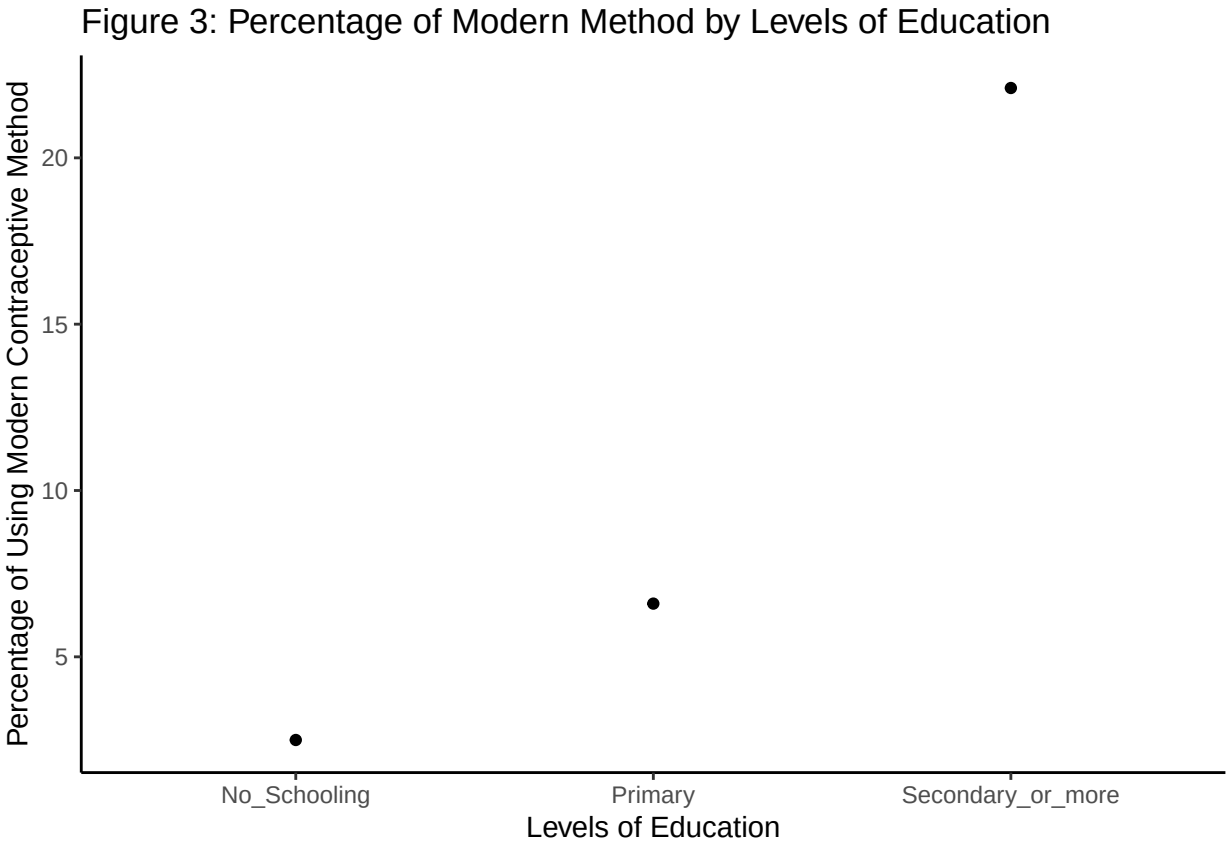


Education

Furthermore, education was the factor that could affect people' decision on using contraception, which requires knowledge for effective birth control. The educational level was a clear criterion of reflecting. By the scatterplot, we could see a very clear trend on how the usage of contraceptive methods varies with respondents' level of education. Individuals that received secondary education or more are obviously more willing to use modern contraceptive methods; over 20% of them used contraceptive methods (Table 3). There was a significant gap between secondary education and other educational levels.

Table 3: Summary By Education

Education	Any_Method	Any_Mordern_Method
No_Schooling	2.8	2.5
Primary	7.3	6.6
Secondary_or_more	26.8	22.1



Hence, from this section, we conclude that educating people that are non- schooling or primary education is the most important part; we need to let them realize the significance of birth control and how to correctly use it.

Tribe

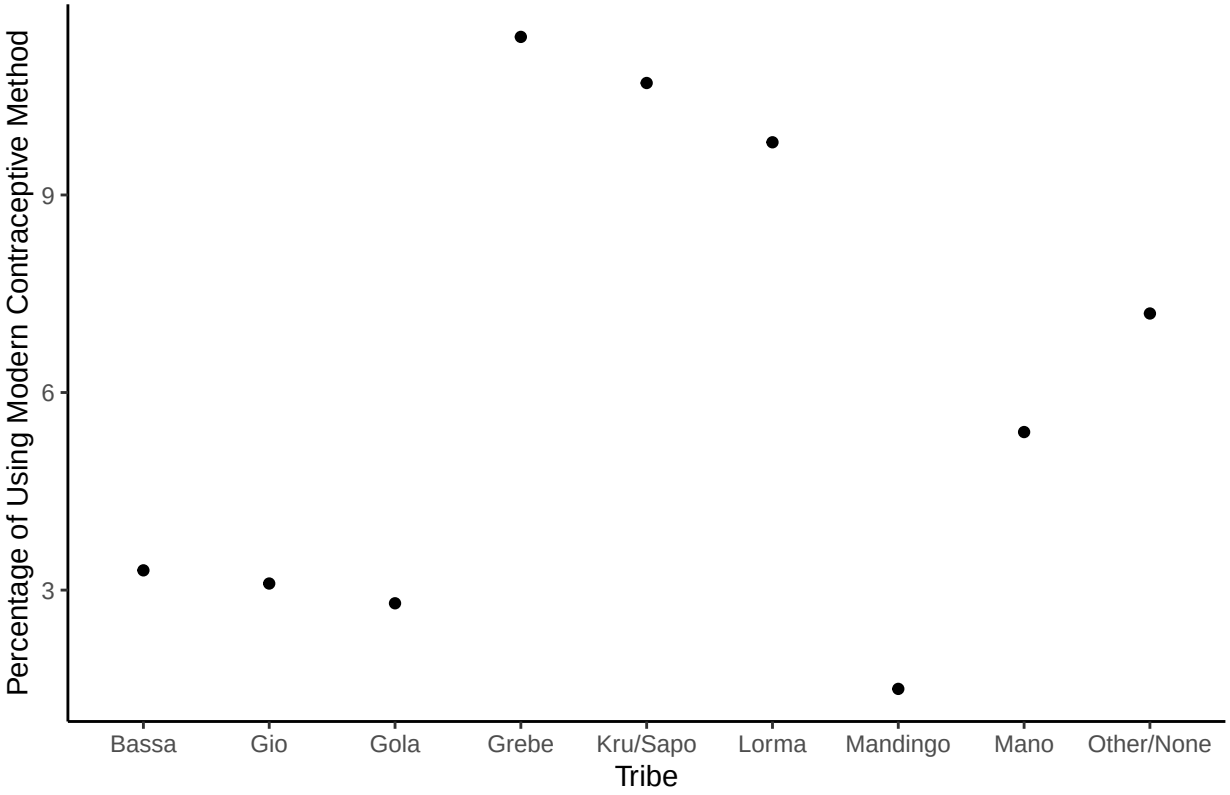
The last factor was a cultural factor. Undoubtedly, the culture has impacts on peoples' way of thinking, which potentially affects their decision on having children. In Liberia, tribes where the respondent belonged was the factor that could identify people from various backgrounds. Liberia has tribes with diverse backgrounds and living habits, we use scatter plots to compare how tribes relate to the use of contraceptive methods. Obviously, tribes truly affect the percentage of women' choices on contraception and there was a large difference between different tribes. "Grebe" and "Kru/Sapo" were the two tribes that have the greatest number of women who choose to do birth control, more than 10% of the respondents choose using contraceptive methods. The tribe "Mandingo" made a large contrast; only 1.5% of the population used contraception (Table 4).

Table 4: Summary By Tribe

Tribe	Any_Method	Any_Mordern_Method
Bassa	4.0	3.3
Gio	3.6	3.1
Gola	2.8	2.8
Grebe	12.2	11.4
Kru/Sapo	11.3	10.7
Mandingo	1.5	1.5
Lorma	12.1	9.8
Mano	6.3	5.4

Tribe	Any_Method	Any_Mordern_Method
Other/None	9.6	7.2

Figure 4: Percentage of Using Modern Contraceptive Method by Tribe



Discussion

The paper examines the percentage of the population that uses contraceptives and the percentage of the population that uses modern contraceptive methods. It is important to examine the amount of population that are still using traditional methods, the difference between the two percentages given in the tables. The purpose is to identify what is the relationship between the use of contraceptives and the groups identified in the paper.

Age

The amount of people using contraceptives for different age groups is on the rise. People used more contraceptives as they aged from 15 to 34. The distribution was bimodal with the peak at 25-29 years olds and 40-44 years olds. Both groups had more than 8% of the population using contraceptives (Figure 1).

Younger individuals, especially people who are below 20, had the lowest rate. We believe that there are two reasons that can potentially explain this. Firstly, teenagers under 20 years old tend to be more less sexually active compared to people who are older for various reasons. That contributes to the total number of people who are in need of contraceptives. Another potential factor is that teenagers have less exposure to sexual education. Their knowledge in engaging in sexual activities safely is limited compared to adults. It could be possible that teenagers are not aware of the importance of using contraceptives. However, the second reasoning is doubtable because 2% among the 2.1% of people who are using contraceptives are using safe modern methods. Only 1% of the population from 15-19 years old are using traditional methods (Table 1). Assuming teenagers are engaged in unsafe sexual activite, the results could indicate that teenagers should

realize the importance of birth control in the future. Policy makers should focus on increasing awareness on the use of effective modern birth controls among teenagers.

The bimodal shape of the distribution could be explained by the fact that people tend to be more active once they age. In their late 20s, it is more likely for someone to have a partner compared to when they were below 20 years old. This might contribute to the total amount of sexually active people in the age group and subsequently people who are using contraceptives. The need to use contraceptives at the beginning of marriage slightly declines as people after marriages, after 35 years old. Even the 35-39 group had a lower rate of using contraceptives compared to its adjacent group, 30-34 and 40-44, only 1% used traditional methods. Moreover, the use of all contraceptives surges again when it comes to over 40 years old (Table 1). The reasoning might be that the risk of giving birth increases as age increases after 40.

Even though we can see a trend in the differences across age groups for using all methods of contraceptives, the trend of percentage of people who are using modern methods in relation to all methods is ambiguous. In other words, there is no clear pattern for people who are still using traditional methods in different age groups.

Number of living children

The number of children one has also contributed to their use of contraceptives. This factor has a negative correlation with the percentage of using contraceptive methods; the more children families had, the larger proportion of them will choose to use contraception. There is no clear difference between groups that used particularly more modern methods.

We believe that the reasoning behind the rising usage of contraceptive methods is associated with people's willingness to have another child. They have enough children already, and they are clearly not pursuing more children. Hence, those families choose to use birth control. Having to raise another child puts more pressure on families time wise and financial wise. The more children a family has, it is less likely that the family would want another one. We assume that families would use more birth control once they have more children. The threshold is at 4 children. The rate of contraceptives among that group peaked at 9.7%, the highest among all groups. However, on the contrary, people who have 5 children use slightly less contraceptives than those who have 4 children (Table 2).

Furthermore, a notable finding in this section is that people who have no children have the lowest usage of contraceptives. Only 2.5% of people use contraceptives in this category. What is surprising is that among 2.5% of people who are using contraceptives, none of them are using traditional methods. They all had modern methods. Whether the contraceptives methods they use were effective has contributed to the fact that they do not have children can be further examined. Even if we cannot conclude on a causal relationship based on our analysis, we can say that people who do not have any children all used modern methods. Policy makers can use the information to better implement education on birth control methods in the country.

Education

The paper further examined the use of birth control methods for each education level in the population. Even though the data did not segmented the groups enough, we can see a big surge in the use of contraceptives from no schooling to primary school, by 4.5%. A bigger jump to people who have higher education from primary school, an 19.5% increase. 26.8% of people who have a secondary or higher level of education used birth control (Table 3). This number is the highest percentage compared to all other percentages shown in the original graph. The education section also has the clearest trend in the change of usage compared with other variables. We can say that higher education level is correlated with more usage of contraceptives.

There are two potential reasons that can explain the high correlation. One is that people who have no schooling background are less aware of the importance of contraceptives. People with more education in school are more exposed to sexual education, usually taught in early middle school or high school. This might also explain the big jump between the usage between primary and secondary school. More laws are imposed to set a mandatory program or session during that age to let people be aware of the importance of sexual sagfety. It is possible that the low rate among people with no schooling experience did not go

through a systematic sexual education program; subsequently limiting knowledge to the existence of birth control. However, this was not to say that the methods they use are less effective or less modern than the other groups. According to the results, 2.5% out of the 2.8% of the population that had no education used modern methods (Table 3). This is to show that among people who use birth control, their methods were scientifically reasonable.

Another reason that can explain higher usage of contraceptives among people with higher education is the opportunity cost of not using contraceptives. Engaging in sexual acts without protection could often lead people to pregnancy. While assuming people with higher education are correlated with higher income, the opportunity cost of bearing a child and raising one would be higher. The income is lost when people are spending time taking care of the child and during pregnancy instead of working. Not using birth control can expose people to the risk of greater loss in opportunities in their career. To avoid the lost opportunity of earning money, more people with high educational backgrounds use more contraceptives than people who have less opportunity costs of having children.

As mentioned earlier in this section, the data could be more segmented into more levels of education to further verify the correlation between usage of birth control and education level. More groups can include university, post-graduate education and other possible educations. Another way of segmenting is that the researchers can specify the years of education the respondents have recieved. It would also be interesting if they asked whether if they had a sexual education class or program in the past. We can then see if the previous potential reasoning about limited systematic sexual education among people with less education experience holds true.

Tribes

Since the data was collected in Sinoe and the Grand Gedeh counties, the cultural difference across different tribes may play a role in the usage of contraceptives. The cultural factors mainly influence people's beliefs about having children and ultimately birth control. The tribes with the most common usage of contraceptives are Grebe, Kru/Sapo and Lorma, all three tribes had an usage of more than 10% of the population. Tribes with lower percentages such as Bassa, Gio, Gola and Mandigo all had a percentage of less than 5%. Among all tribes, Grebe had the highest usage and Mandigo had the lowest (Table 4).

By examining the percentage of the contraceptive methods that are considered modern, all people from Gola and Mandigo who use birth control use modern methods. No one in the two tribes are using traditional methods. They are also two tribes with the lowest usage of contraceptives in general. This indicates that the culture might be influencing how using contraceptives is common, but the method they use when they decide to use it is effective.

To further research about the reasoning behind the difference in rates among the tribes, future research can provide more background information on the difference in belief among the tribes. They can see the attitude of people from different tribes about having children and having sex. They can also look at the average number of children per household, age of first pregnancy, age of marriage as qualitative factors. For the limitation of this paper, we only focused on the difference between tribes on whether they use modern methods.

Limitations

For the data collected, only numerical values provided are the percentages. Limited analysis can be done on two columns of percentages where one column is included in another. The 'any method' column includes both modern and traditional contraceptives methods. The only analysis we can perform on the data is observing differences across different characteristic groups such as age, education, etc. Moreover, we cannot examine the characteristics of the people in one category. For instance, we might be interested in seeing how many people within the category of having two kids who have no education backgrounds. Analyzing from an organized table as used in the paper as opposed to starting from the raw dataset can very much limit potential findings.

Another limitation with the data provided is that the table used in the paper does not provide a full picture. The percentages shown only provide the proportion of the entire population that uses contraceptives. Without

any context, we do not know what percentage of the population are sexually active. There exists a gap where we cannot conclude on what percentage of the people who are active are using contraceptives. We can only observe the difference in the methods used among that group of people who use contraceptives.

In future research, the paper can examines the data with a broader view of the situation and connect the difference in contraceptive methods with other variables such as the percentage of the population that are sexually active, marrital status, beliefs, and other potential factors. Despite the lack of context in the research, the time when the survey was taken was 1984. The analysis can only be used to see what was the situation at that time. It cannot be representative of the population today since more methods have been developed and people's beliefs have changed drastically. Perhaps another topic to research in the future is about how people's use of contraceptives change over time compared to today's data.

Appendix

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 paper was created to present the findings of the Liberia Demographic and Health Survey. It was conducted under a program called Demographic and Health Surveys, designed to collect data on fertility and family planning.

2. *Who created the dataset (for example, which team, research group) and on behalf of which entity (for example, company, institution, organization)?*
    - The research was conducted by the Ministry of Planning and Economic Affairs in 1987. The survey was carried out with the assistance of the Institute for Resource Development.

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 research has tables, figures and a map that shows and supports the findings. In additional to those, the paper also provided a background section in text, providing context for the findings. The survey questionnaires are attached in the appendix.
2. *How many instances are there in total (of each type, if appropriate)?*
    - There are 53 tables, 8 figures and 1 map.
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 paper contains all tables, figures and maps. The survey was conducted to the general population of Liberia, it is representative of the whole country.
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.*
    - All data from the questionnaires are entered on to microcomputers. The tables and figures we see in the paper are all processed and properly laid out by sections.
5. *Is there a label or target associated with each instance? If so, please provide a description.*
    - Yes. The paper was divided based on general topics. THe topics include: Marriage and Exposure to the risk of pregnancy, fertility, contraceptive knowledge and use, fertility preferences and mortality and health. Each instance is displayed under the corresponding sections with titles.
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.*
    - No
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.*
    - Yes, the organization of the paper made the connection 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.*
    - No
9. *Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.*
    - The results are affected by non-sampling error and sampling error. The sampling method used was not simple random. The survey had a standard error. The non-sampling error is due to the mistakes made in carrying out field activities, such as failure to locate and interview the correct

household, the way questionnaires were sked, misunderstanding of the questions etc. Non-sampling errors are impossible to eliminate.

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.*
    - It is self-contained. The paper was based on the dataset they collected.
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.*
    - No
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.*
    - No
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 research only surveyed women in Liberia.
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.*
    - No
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.*
    - 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 was collected based on survey responses of 5026 households interviewed. The raw data consisted of 5026 observations. The results were organized into the paper. The data was not inferred or derived from other 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 field staff for the LDHS consisted of 24 female interviewers, 6 field editors, 6 supervisors, and one fieldwork coordinator.6 teams are formed based on major dialects to survey the entire sample. After interviews, data from the questionnaires were entered onto microcomputers at the Bureau of Statistics office in Monrovia. The data were then subjected to extensive checks for consistency and accuracy.
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 survey was a national-survey. Every women in the country was given the survey. The response rate was 90%.
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 field staff from the survey program was involved. 24 female interviewers, 6 field editors, 6

supervisors and 1 fieldwork coordinator. The paper did not mention the compensation.

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 collection began in late February and largely completed by July 1986.
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.*
   - No
7. *Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (for example, websites)?*
   - The data was obtained directly from the respondents.
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.*
   - Yes, they are aware of the collection. At the top of the questionnaire, there was 'I would like some information about the people who usually live in your household or are staying with you now.'
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.*
   - By responding, they consent to give their information.
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).*
    - No
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.*
    - 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 processing process was not mentioned in the paper.
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.*
   - N/A
3. *Is the software that was used to preprocess/clean/label the data available? If so, please provide a link or other access point.*
   - N/A
4. *Any other comments?*
   - N/A

Uses

1. *Has the dataset been used for any tasks already? If so, please provide a description.*
   - No
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.*
   - No
3. *What (other) tasks could the dataset be used for?*
   - The survey will also be useful in providing baseline information to the Southeast Region Primary Health Care Project (SER/PHC) for planning its program to motivate and educate the people in Sinoe and Grand Gedeh Counties in preventive health measures.

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?*
    - No
5. *Are there tasks for which the dataset should not be used? If so, please provide a description.*
    - 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.*
    - No
2. *How will the dataset be distributed (for example, tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?*
    - The paper did not mention how the dataset will be distributed. The paper is accessible to the general public in many databases.
3. *When will the dataset be distributed?*
    - The paper did not mention when the dataset will be distributed.
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.*
    - N/A
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
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.*
    - No

Maintenance

1. *Who will be supporting/hosting/maintaining the dataset?*
    - The survey was conducted by the Demographic and Health Surveys Program. The paper did not mention if they are going to conduct future surveys.
2. *How can the owner/curator/manager of the dataset be contacted (for example, email address)?*
    - Demographic and Health Surveys Program can be reached through https://dhsprogram.com/Who-We-Are/Contact-Us.cfm.
3. *Is there an erratum? If so, please provide a link or other access point.*
    - No
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)?*
    - N/A
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.*
    - No
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.*

- No

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.*

- They would have to contact the Demographic and Health Surveys Program.

Reference

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Franbois, Yihui Xiein, Lionel Henry, and Kirill Miller. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*.

Henry, Lionel, and Hadley Wickham. 2020. *Purrr: Functional Programming Tools*. https://CRAN.R-project.org/package=purrr.

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/.

Statistics, Bureau of, and Ministry of Planning. 1986. *Demographic and Health Survey*. https://dhsprogram.com/pubs/pdf/FR22/FR22.pdf.

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

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain Franbois, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.%0A%20%20%20%2021105/joss.01686.

Wickham, Hadley, Romain Franbois, Lionel Henry, and Kirill Miller. 2020. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.