Winners and Losers: Examining School Enrollment Rates in

Post-Civil War Liberia

Emma Mayfield

Advisor: Professor Paul Cichello



Boston College

Department of Economics

Senior Honors Thesis

May 2023

Acknowledgments

Thank you to my thesis advisor, Professor Cichello for making this thesis a reality. It would not have been possible without your support and invaluable guidance. Not only because of your help throughout the process, but also because your courses sparked my interest in econometric techniques and development economics, and ultimately led to my decision to pursue a thesis in the first place.

Thank you to Professor Murphy and Professor Grubb for guiding me through the thesis process. I also want to thank all of the professors I have had at BC, particularly in the Economics and International Studies departments. Every one of you has had a unique impact on me and my education.

Thank you to my wonderful friends and roommates, Holly and Alyssa, who put up with me talking through my ideas and complaining about issues I ran into for an entire year, even though I am sure they never want to hear about school enrollment in Liberia ever again.

Lastly, thank you to my family, specifically my parents, for your constant love and support. You've given me so many opportunities, including going to Boston College, and you made me the person I am today. I love you and I owe all my successes to you.

Table of Contents

Abstract1
Introduction
Background4
Literature Review
Bounce-back17
Data22
Dependent Variable
Age23
Sex and Location Variables
Wealth Index27
Education Loss27
Prewar Education35
Predicted Education Gain41
Conflict42
Aid48
Bounce-back49
Change Over Time51
Methods
Linear Probability Models62
Probit Models64
Results

Was there bounce-back?	83
Robustness Checks	85
Overage Students	
Using Different Ranges for Education Loss and Gain Variables	91
Conclusion	95
References	99

Abstract

Liberia had two devastating civil wars 1989-2003. I am examining who benefitted from the large amounts of international aid and development programs that poured into the country during the post-war rebuilding period, in terms of school enrollment rates. With USAID's Demographic and Health Surveys and Uppsala Conflict Data Program's Georeferenced Event Dataset, I use probit and linear probability models to examine the determinants of being enrolled in school in 2007 and 2019. I find that females and kids living in rural areas had disproportionate recovery in the post-war period controlling for other explanatory variables. Household wealth was an important factor in determining enrollment. I also examine the concept of bounce-back, or rapid recovery in post-conflict contexts. I find that on a national level, there was significant recovery in enrollment rates, with about 51% of kids being enrolled in school in 2007 and about 81% being enrolled in 2019. I was unable to determine definitively whether or not this recovery was proportional to the amount of loss experienced due to the wars due to large standard deviations.

Introduction

Liberia had two civil wars in 1989-2003, which were devastating to the country in both physical destruction and lives lost. After the civil wars, Liberia had unprecedented levels of aid and recovery programs pour in from various countries and inter-governmental organizations. I am examining who, if anyone, benefitted the most from this aid and these programs in terms of school enrollment in the post-civil war period. I will be using USAID's Demographic and Health Surveys for Liberia in 2007 and 2019 for enrollment data and all of my explanatory variables except conflict. For my conflict data, I will be using Uppsala Conflict Data Program's Georeferenced Event Dataset.

To determine what groups saw the most improvement in enrollment levels, I will first be looking at the change over time in enrollment levels for 6-18 year-olds. I found that overall enrollment rates improved from 51.01% in 2007 to 81.04% in 2019. Females saw more improvement than males, rural areas saw more improvement than urban areas, and areas with lower level of conflict during the wars saw more improvement than areas with higher levels of conflict. There was fairly consistent improvement in enrollment rates across different levels of household wealth.

To examine which factors were the most important in determining the likelihood of a given kid being enrolled in school, I will be using probit models and linear probability models for 6-18 year-olds. My dependent variable is a dummy variable that equals 1 if an individual is enrolled in school. The explanatory variables I will be using are sex, age, age squared, urban or rural area, household wealth, conflict levels in each area during the wars, levels of education in each area before the wars, how many years of education each area was predicted to gain during the war period had the war not occurred, and how many years of education people on average in

each area lost because of the war. I found that female and urban were significant indicators of the probability of being enrolled in 2007, but not in 2019, suggesting that, during the rebuilding period, females and people living in rural areas caught up to males and people living in urban areas controlling for other explanatory variables. Household wealth was an important indicator of the probability of being enrolled in both years. I used deaths due to the wars as a measure of conflict and found that number of deaths in a certain area was not an important indicator of the probability of being enrolled in 2007, but in 2019, more deaths in an area was associated with a lower probability of being enrolled, suggesting that loss of life in a certain area had a long-term impact on the probability of kids in that area being enrolled in school. The levels of prewar education in a given area were significant indicators of the probability of being enrolled in both 2007.

Bounce-back is the idea that post-conflict societies will have significantly higher levels of growth immediately after the conflict than societies that have not experienced conflict. These growth levels allow societies to recover or 'bounce-back' from the conflict. I am expanding this idea to enrollment rates to examine whether or not there was bounce-back in enrollment rates in Liberia in their rebuilding period. Because I do not have data for enrollment before the war, rather than looking at overall recovery on a macro-level, I am looking at relative recovery in Liberia. I am using education loss due to the war to examine whether or not people with more loss experienced proportional recovery to people with less loss. Due to large standard deviations, the model was not able to determine concretely whether or not this proportional recovery or bounce-back occurred, but the best guess given the sample is that there was proportional or near-proportional recovery. However, I could not rule out the possibility that education loss had a positive or negative effect on enrollment.

Background

In 1822, former slaves and free black people from the United States established the colony of Liberia. There were already people living in the area that would become Liberia, and these indigenous Liberians were largely cut out of the political process of the new Liberian government. This divide between the elite Americo-Liberian minority which ran the government and the indigenous Liberian minority would set the scene for decades of political turmoil.

Firestone established the first rubber plantation in Liberia in the 1920s, which began one of Liberia's most profitable and important industries. World War II greatly increased the demand for rubber¹ and iron, another of the Liberia's main exports. These increased exports greatly grew the Liberian economy. This growth continued fairly steadily through 1971 under President William Tubman, a member of the Americo-Liberian elite. Tubman was replaced by William Tolbert, his second in command. Growth continued in the 1970s but did begin to slow due to global stagnation and a decrease in mining exports in 1975 after a decrease in the price of iron.

This economic growth, however, was not equally enjoyed by all Liberians. The Americo-Liberian elites saw much of the benefit, and wealth inequality continued to grow between these elites and the indigenous Liberians. The growth under Tubman and Tolbert was criticized as 'growth without development'². Tension began to grow in response to this inequality in the 1970s, with the non-elite demanding more political power and the elites unwilling to give up any of their influence over the government.

In 1979, the government spent millions of dollars hosting a conference for the Organization of African Unity, putting the country further in debt.³ The government then

¹ (*Doss* 2020)

² (Werker and Beganovic 2011, p.3)

³ (Werker and Beganovic 2011)

increased the price of rice, a staple for the majority of Liberians. The price hike led to demonstrations and riots, which quickly turned violent. The violent demonstrations combined with the growing political tension, and eventually led to a coup in 1980. Samuel Doe, an indigenous Liberian, and his supporters assassinated Tolbert and took control of the government, promising an end to Americo-Liberian dominance.⁴

With a sudden overthrow of the established system in Liberia and many of the elite fleeing or being displaced, the economy quicky stagnated, as seen in Figure 1.



*Figure 1: Graph of GDP in Liberia*⁵

The economic stagnation only exacerbated the tensions stirred by Doe's coup. There was another coup attempt in 1985 against Doe.⁶ The attempt failed, but in response Doe launched a violent campaign against the Gio and Mono ethnic groups in central Liberia, because large portions of those groups had supported the coup.⁷ The United States government had previously

⁴ (Doss 2020)

⁵ (FRED)

⁶ (Werker and Beganovic 2011)

⁷ (*Doss* 2020)

been giving the Liberian government hundreds of millions of dollars in aid, but they stopped in the hopes that the Doe government would be replaced.

With the repression and economic instability of the 1980s, a larger conflict soon broke out. In 1989, Charles Taylor led the National Patriotic Front of Liberia (NPFL) in an uprising against Doe, assassinating him in 1990. This started a 7-year, extremely brutal civil war. The civilians in Liberia bore the brunt of the devastation, with as many as 250,000 murdered and more than a million fleeing into exile or being displaced.⁸ This is a huge amount of disruption, considering the population of Liberia was only 2.5 million people in 1989.⁹

The fighting came to an end in 1997 with the overthrow of the existing government and election of Taylor to the presidency. Taylor was a brutal leader who was guilty of several war crimes during the first civil war. While there was little fighting, his government also provided almost no public services, and public functions like schools were not running at anywhere near their previous capacity. ¹⁰

In 2000, a new rebel group, Liberians United for Reconciliation and Democracy (LURD) formed to oppose Taylor. Another civil war broke out between Taylor's government, LURD, and Movement for Democracy in Liberia (MODEL), an anti-Taylor rebel group backed by the Cote d'Ivoire government.¹¹ The second civil war ended in 2003, with Taylor fleeing to Nigeria.¹² An agreement for a UN-backed ceasefire was reached between the rebel groups to form a democratic government free of Taylor. President Ellen Johnson-Sirleaf was elected in 2005.¹³

⁸ (International Peace Institute 2008)

⁹ (World Bank)

¹⁰ (International Peace Institute 2008)

¹¹ (*Reuters* 2011)

¹² (International Peace Institute 2008)

¹³ (*Reuters* 2011)

Liberia remained relatively stable after the ceasefire, with only minor outbreaks of violence; however, the 14 years of civil war had devastating effects on the country.¹⁴ Before the wars, Liberia was rapidly approaching the middle-income threshold, with an average annual economic growth rate of 7% from 1955-1975.¹⁵ The civil wars destroyed up to 90% of the country's economy.¹⁶ The wars left approximately 340,000 children, 18% of children in Liberia, orphaned. There were approximately 21,000 child soldiers in Liberia by 2003.¹⁷

After Johnson-Sirleaf's election and the UN lifted the diamond embargo in 2007¹⁸, foreign aid started flooding into Liberia (Figure 2).



Figure 2: Estimated Foreign Aid Received by Liberia 2005-2016¹⁹

- ¹⁵ (Waydon, Ying, and Ketter 2016)
- ¹⁶ (*The Economist* 2017)
- ¹⁷ (Bosede 2012)
- ¹⁸ (Reuters 2011)
- ¹⁹ (*The Economist* 2017)

¹⁴ (*The Economist* 2017)

The government created an Aid Management Unit within the government to manage the incoming foreign aid.²⁰ The goal of the government agency was twofold. First of all, they wanted to maximize the effectiveness of the aid. Funneling the aid through a central agency lessens the risk of overlapping or contradictory programs.²¹ Second of all, the government was new and wanted to maintain legitimacy in the eyes of their citizens. It is much easier to gain their support if the programs and new infrastructure are coming through the government (even if not funded by the government) rather than coming straight from foreign backers. However, due to concerns about waste and corruption in the new government, much of the aid was 'off-budget', and did not go through the government, making it difficult to keep track of the hundreds of millions of dollars pouring into Liberia in the years following the civil wars. In the 2009-2010 fiscal year, there was \$23 million of foreign aid in the official Liberian budget, but there was an estimated \$449 million of off-budget foreign aid. Several of the major donors only used off-budget foreign aid including the United States and most of the UN agencies.²²

The sheer number of donors also made it incredibly difficult to keep track of the amount of aid flooding in. Several governments from around the world contributed, including the United States, the United Kingdom, the European Union, Sweden, Ireland, Norway, Japan, and China. There were also many inter-government organizations getting involved, such as the World Bank, UNICEF, the World Health Organization, and many United Nations agencies.²³ This is not even taking into account the smaller NGOs that funded projects or the projects which spanned multiple countries. For example, the Gates Foundation funded a project aimed at rural famers in

²⁰ (Parks and Kadaba 2010)

²¹ (*The Economist* 2017)

²² (Parks and Kadaba 2010)

²³ ("Donor Coordination" 2022)

eight different West African countries including Liberia. It is not clear how much of the \$90 million of funding went into Liberia specifically.²⁴

Liberia recovered well economically following the civil wars. The economy steadily improved until a stagnation around 2014 because of an Ebola outbreak, and a dip around 2020 because of the Covid-19 pandemic.





While it is fairly clear that the economy recovered in Liberia on a country-level at least in part due to aid, it is much less clear how outcomes changed for individuals in Liberia in the years following the wars. The individual outcomes do not necessarily reflect the country-level outcomes, especially when aid is concerned. A good example of this is the China Union takeover

²⁴ ("Gates Foundation, Partners Pledge \$90 Million to Boost Incomes of Small Farmers in Africa" 2009)

²⁵ ("GDP per Capita (Current US\$) - Liberia | Data" n.d.)

of the Bong Mines. In 2009, the Liberian government signed a 25-year, \$2.6 billion contract with China for the right to control and rebuild the Bong Mines, which were iron ore mines destroyed during the civil wars. On a country-level, the deal was a huge win for Liberia. China Union hired over 3,000 Liberian workers and Liberia started seeing royalty returns within the first couple years of operation. China also planned to build new roads, schools, hospitals and a hydroelectric plant in the area. However, the deal seemed to be far less favorable for the individual Liberians actually effected by the project. In 2013, the workers at the mines went on strike because of the abusive labor practices and low pay. Reports coming out of the mines described the conditions as akin to "modern slavery". The people living in the areas surrounding the mines complained of a lack of basic services and contaminated water.²⁶

Different people benefitted from aid and rebuilding to varying degrees because the aid programs were not distributed evenly across the country. For example, USAID's programing in Liberia focused on six of the fifteen counties: Bong, Grand Bassa, Lofa, Margibi, Montserrado, and Nimba.²⁷ However, there has been little research done on what individuals benefitted from rebuilding process, as most of the analysis is focused on the country as a whole. Because of that, I want to focus my analysis on individuals to see what types of people benefitted from the rebuilding process.

Literature Review

Recovery in post-conflict areas is based on several factors. The most obvious factor is aid. However, there are countless other variables at play such as strength of institutions, security

²⁶ (Roessler and Blair 2014)

²⁷ ("Education | Fact Sheets | Liberia | U.S. Agency for International Development" 2022)

and stability, and resilience of the community. Economies, even without an abnormal influx of aid, will experience some degree of 'bounce-back' in the years following a conflict. Their economy has been greatly suppressed, or even ruined, in a major conflict, so there is significant opportunity for growth as the infrastructure begins being rebuilt, people go back to work, and normal economic activities restart.

In a study of 17 countries post-civil war, Collier and Hoeffler examine their growth relative to the standard growth of countries that did not experience a civil war or any major conflict. They found that, controlling for policies, institutions, governance, and aid, post-conflict countries grew on average 1.13% more rapidly than non-conflict countries. In addition, post-conflict deviations from normal growth typically followed an inverted u shape, with the countries remaining relatively stagnant in the immediate aftermath of the war, then experiencing rapid growth as rebuilding begins to pick up and everything begins to return to prewar conditions, and then eventually tapering off to a normal level of growth as the development converges with prewar levels of development. This peak phase for growth is typically 4-7 years after the end of the conflict.²⁸

²⁸ (Collier and Hoeffler 2004)



*Figure 4: Liberia Growth Rate (%) 2001-2021*²⁹

Based on Liberia's growth rate, Collier and Hoeffler's trend holds. The growth rate reached a peak of 9.5% in 2007, four years after the end of the violence in 2003. Although, Liberia did maintain a high level of growth (between 5.3% and 9.5%) between 2005 and 2013, which is longer than a high level of growth is typically maintained according to Collier and Hoeffler's study. 2014 was also the year of a major Ebola outbreak in Liberia, which likely accounts for a lot of the drop off in growth (Figure 4).

Where Liberia differs from the cases in Collier and Hoeffler's study is in regard to aid. First of all, they found that aid is not substantially higher in post-conflict countries than it is in similar countries that did not experience conflict.³⁰ This is not the case in Liberia. Liberia is a

²⁹ ("GDP Growth (Annual %) - Liberia | Data" n.d.)

³⁰ (Collier and Hoeffler 2004)

relatively small country, with a population of around 3,000,000 directly after the war in 2003.³¹ It was also heavily devastated, more than is typical of this type of conflict, during its civil wars. Because of this, it was a good candidate for international aid. Countries and organizations could come in and build new infrastructure or start programs and instantly see the impact of their funds.³² Aid poured into Liberia at a much higher rate than it did into similar no-conflict countries during the same time period.³³

Collier and Hoeffler also found that when there is greater aid for post-conflict countries, the aid usually flows into a country directly after a conflict, decreases substantially after the 2nd year of post-conflict, and then is back to its normal levels six years after the end of the conflict.³⁴ As previously discussed, this is not the case with Liberia as aid began coming in great numbers four years after the conflict in 2007, then decreased 7 years after the conflict in 2010, but still maintained fairly substantial levels after 2010.³⁵ The aid also coincides with peak growth in Liberia. Is the recovery in Liberia due to the aid, the natural growth peak demonstrated by Collier and Hoeffler, or a combination of the two?

In addition to the economy as a whole, conflict, and civil wars specifically, significantly harm the education system and enrollment in the countries in which they occur. In a study of countries 1980-1997, Lai and Thyne found that on average, countries in civil wars decrease school expenditure by 3.1%-3.6% each year. They also experience on average a 1.6%-3.2% decrease in enrollment with the biggest decrease in tertiary enrollment, then secondary enrollment. The severity of the civil war is also important. They

- ³³ ("Total Official Development Financing ODF | OECD International Development Statistics | OECD ILibrary" n.d.)
- ³⁴ (Collier and Hoeffler 2004)

³¹ ("Population, Total - Liberia" 2023)

³² (*The Economist* 2017)

³⁵ (*The Economist* 2017)

estimate that for every additional 1,000 people killed yearly in a civil war, school expenditures decrease 2%-2.7% and enrollment decreases 1.4%-3.4%. They also find some evidence to suggest that male tertiary enrollment typically decreases more than female tertiary enrollment, because more young males are recruited, willingly or unwilling, to join the fighting.

Lai and Thyne hypothesize that these decreases in enrollment specifically are due to the displacement of students, the destruction of facilities, and in some cases, the deaths of school aged children. They point out that in Liberia, 80% of schools were closed at some point during the first civil war (1989-1997). That fact combined with the high degree of destruction and lethality of the civil wars in Liberia would suggest that the decreases in both school expenditure and enrollment during the civil wars were likely very large.³⁶

A 2007 study from Liu confirms and expands this hypothesis. Liu looks at education outcomes for Liberians who were school age during the wars and discusses how those outcomes affect long-term economic prospects and political participation compared to the outcomes of people who were kids before the war. Using a difference-in-differences model where the control was areas with low education levels before the war and the treatment was areas with high education levels, she found that in areas of high education, people who were children during the wars had significantly less education than the people in the same areas before the war. She points out that during the war not only were some students and teachers shot in schools, but a lot of the schools were also bombed or taken over by the rebels to be used as bases, leading to an obvious decrease in enrollment.

These decreases in educational opportunities had long-term consequences for the people who were children during the war. It is not common for people to go back and get a basic

³⁶ (Lai and Thyne 2007)

education once they are adults. Many children who were still of school age directly after the war could not return to school afterwards because the economic loss from the wars meant that they could not afford the fees, or their families needed them to work. It also took a little while to rebuild infrastructure and get the school system up and running again. This meant that people who lost educational opportunities generally also lost long-term economic prospects, because they had to join the unskilled labor force.

Liu highlights that the loss of educational outcomes is fundamentally different than a lack of education in general when it comes to individuals' attitudes and how they view their own outcomes. People from areas with low education before the war did not have their educational opportunities affected as much because of the war because they were unlikely to have as many educational opportunities relative to the rest of the country had the war not occurred. They did not have as many expectations about having certain educational opportunities. The general destruction harmed their future economic prospects because it destroyed the economy as a whole and many job opportunities, but likely did not largely change their educational outcomes. Many of the people from low-education areas always expected to join the unskilled labor market. After the war, the majority of new jobs were manual labor jobs aimed at rebuilding infrastructure. So, while the war was disruptive and no doubt traumatic for people in areas with lower initial education levels, it did not harm their long-term economic prospects in the same way as people who expected to have higher paying jobs but lost the opportunities they likely would have had if the war had not occurred. People who expected to have relatively high incomes had outcomes that were much worse than their expectations and this mismatch of expectations and reality led to far greater discontent with the government and lower political participation among the people who came from areas with relatively high educational outcomes before the war.³⁷

Liberia and its donor partners have made a concerted effort to improve the education system and make it more equitable since the wars. Liberia made primary education free and compulsory in 2001, and reaffirmed the system in 2011, expanding it to reflect the realities of a country no longer in the midst of a war.³⁸ They set several goals to improve both quality and quantity in the education system. The school feeding program, in which they provided lunches for students who could not afford to bring their own and sent food home with especially needy children was especially instrumental in improving retention rates.³⁹

In 2005, the Ministry of Education created an education policy specifically aimed at making education more equitable and available for girls in Liberia. They established a Girls' Education Unit within the Ministry of Education and set a goal that by 2015 girls would have equal access to education and there would be an equitable geographic distribution. The plan to accomplish this was to create a nationwide education campaign about the importance of female education, take steps to protect girls from sexual assault in school and on their way to and from school, and recruit more female teachers and administrators.⁴⁰

The results of these efforts are mixed and highly debated. There is some indication that the probability of attending public education has risen, indicating that more children are taking advantage of the free primary school. However, there are still major costs associated with even public education.⁴¹ In addition in Montserrado, the most populous county, and Margibi county,

³⁷ (Liu 2022)

³⁸ (Waydon, Ying, and Ketter 2016)

³⁹ ("Liberia - Joint IDA-IMF Staff Advisory Note on the Poverty Reduction Strategy Paper : First Annual Progress Report" 2010)

⁴⁰ ("Policy Brief: Liberia Girls' Education Policy" n.d.)

⁴¹ (Cuesta and Abras 2013)

there are more private than public schools. This means that in the main city in Liberia it is more difficult for children to afford school. In addition, the education system is still highly decentralized, with the local governments running the systems in their areas with little oversight, leading to lax enforcement of government initiatives⁴² and separate NGOs running different education programs and curriculums in different areas.⁴³

Bounce-back

Bounce-back is a concept associated with areas recovering or 'bouncing back' after a conflict. The basic idea behind bounce-back is that a country post-conflict will grow at a higher rate than that same country would grow had there not been a conflict. This is because damage to the country's infrastructure and economy as a whole allows for rapid growth of physical capital as the country rebuilds itself back to its pre-conflict level. Most of the human capital remains in the country (besides the people who die or become refugees), so at the end of a war there are generally large amounts of untapped human capital unable to be utilized due to the disproportionate lack of physical capital. However, after the war, people go back to work, infrastructure is rebuilt, and normal economic activities restart, rapidly growing the economy from its suppressed state during the war. Data shows that post-conflict countries generally have more economic growth than countries that have not had a conflict.⁴⁴

⁴² (Waydon, Ying, and Ketter 2016)

⁴³ (*The Economist* 2017)

⁴⁴ (Collier and Hoeffler 2004)





Figure 5 above visualizes this concept. This is an example that does not use data from an actual country. The normal growth rate for an economy varies greatly from country to county, but in this example, this is a county with a fairly steady positive growth rate. During conflicts, economic activities are greatly suppressed, and infrastructure is destroyed, commonly causing economies to actually have negative growth during a conflict. This is shown in the red 'Conflict' time period. The growth rate quickly falls to a negative level and stays at a negative level throughout the conflict. After the conflict ends, the growth rate is even more positive than it was pre-conflict as the county rebuilds. This is the 'Bounce-back' section of the graph. The level of growth should theoretically decrease to around the same level as before the conflict as exogenous growth related to the conflict, like infrastructure rebuilding, ends, and the country goes back to the level of economic activities that was normal during the pre-conflict period. This return to normal is visualized in the 'Post-Conflict' section of the graph.

The more devastating and disruptive a conflict is, the greater the amount of bounce-back an area will experience. For example, if everyone in an area stops working, the economic growth will be worse off than if half of the people stop working. Similarly, if everyone was previously out of work and returns at the same time, then the growth rate will be higher than if half the population continued to work during the conflict and half came back at the same time. More disruptive conflicts have a lower absolute starting point at the end of a conflict, so more room for rapid growth.



Figure 6: High vs Low Destruction Bounce-back

Figure 6 demonstrates the idea of low vs high destruction/disruption conflicts. Generally, the more destruction, the more room for growth.

Theoretically, this should have also happened with enrollment rates. Students were physically unable to go to school during the war, which would have greatly lowered enrollment rates. 80% of schools in Liberia were closed at some point during the first civil war between 1989 and 1997.⁴⁵ However, there were still large amounts of students (human capital) in Liberia with a demand for schooling, so there was a mismatch of demand for schooling and supply of schools. As school are rebuilt and reopen and daily life begins to go back to normal, enrollment levels should go back to near prewar levels as the physical capital recovers to the level of human capital.

Groups with higher levels of enrollment before the war should theoretically have greater levels of bounce-back, because the sudden loss in enrollment would have been more 'destructive' for them. There would have been higher levels of loss for groups that started at a higher level of enrollment, so there should also be more room for bounce-back for those groups. Groups with lower levels of enrollment before the war did not experience as much loss (as many people not enrolled who would be enrolled if not for the war), so they have a lower ceiling for bouncing back to pre-conflict levels.

Boys have historically had higher rates of school enrollment than girls in Liberia, so the bounce-back of enrollment rates by gender should look something like Figure 7.

⁴⁵ (Lai and Thyne 2007)





Note: This graph measures enrollment as proportion of kids enrolled in school in real terms, not in growth rates like the economic growth graphs above. This graph is also theoretical, it does not use real data from Liberia.

The enrollment rates for boys started off higher than for girls up until the beginning of the war in 1989. During the war, enrollment rates for both groups should have decreased due to the school closures and instability, but the rates for boys should have decreased more significantly, as there were a higher proportion of boys already in school to drop out. As the country rebuilds in the post-war period between 2007 and 2019, the same types of people who were enrolled before the war, and in some cases the same people, should return to school. Taking into consideration only bounce-back, the levels of enrollment for different groups should return to approximately the same levels of enrollment from before the war, so boys should have significantly more recovery than girls. Therefore, boys should experience more 'bounce-back'

due to higher initial enrollment levels. This is of course complicated in reality by various other factors like recovery aid and government policies.

Data

The main datasets I will be using are USAID's Demographic and Health Surveys for Liberia in 2007 and 2019. 2007 is the earliest year after the end of the conflict that the DHS collected data. It is also the first year when aid began pouring into Liberia at high levels, and the economy had begun to improve. There are 37,670 observations total. I will be looking at enrollment rates for 6-18 year-olds. There are 11,466 observations for 6-18 year-olds. There are no missing observations for any of the variables I am using in my regression models, so there are also 11,466 observations in the regression models.

2019 is the most recent year that data is available. The enrollment rates from 2019 will show some of the longer-term impacts of the war and the rebuilding process. There are 41,999 observations total in 2019 and 14,633 observations for 6-18 year-olds. There are also no missing observations for any of the variables I am using in the regression models in 2019, so there are 14,633 observations in the regression models.

There were 300 clusters selected in 2007 out of 4,602 possible enumerated areas, and 325 clusters selected in 2019. Sampling weights were calculated for each household from the probability of selection for each household, adjusted for nonresponse. The household weights were multiplied by 1,000,000 for data storage purposes, so I divided them by 10,000 so that the weights accurately reflect the population size of Liberia. Throughout my analysis I uses these weights and Stata's svyset command to accurately reflect the population in strata and primary sampling units.

Dependent variable:

My dependent variable is enrollment. This is a dummy variable that equals one if the person is currently enrolled in school and zero if they are not. Enrollment information was collected for every person aged 3-24 in 2007 and 5-24 in 2019.

Age:

There are observations for people aged 0-97 in the DHS dataset. For the enrollment regressions I will be using data from people aged 6 to 18. 6-18 are the official years of schooling in Liberia. Primary school is ages 6-12.⁴⁶ Despite the official schooling ages many kids in Liberia are overage for their official schoolyear.⁴⁷ This was especially true in 2007, as seen in Figure 8.

⁴⁶ ("Liberia National Education Profile 2018 Update" 2018)

⁴⁷ (The International Bank for Reconstruction and Development and The World Bank 2010)



Figure 8: 2007 Graph of the Proportion of People Currently Enrolled in School by Age

Enrollment did not greatly drop off after age 18. This is likely because many people were not able to go to school during the war, so they returned to school after the war ended, even though they were no longer the official school ages. This trend is essentially a result people starting school late due to the war. The 2019 enrollment data confirms that this late enrollment was mainly due to the war.



Figure 9: 2019 Graph of the Proportion of People Currently Enrolled in School by Age

Enrollment greatly decreases in the upper teens. While some people are still overage for their school year in 2019, it is nowhere near the levels in 2007, which suggests that the 2007 trends were due to people making up education lost during the war.

I considered including people aged up to 24 in the regressions measuring enrollment because there are so many people older than 18 still enrolled, especially in 2007; however, I believe that if I include people aged 19-24, I will also be capturing people who have already completed their education. People who have already completed their education are not enrolled, so the results could be biased. For example, if extremely wealthy people generally completed their education on time and less wealthy people are overage for their school year, wealthy people will be not enrolled in the regression and less wealthy people will be enrolled, suggesting that more wealth makes people less likely to be in school, which is highly unlikely. For this reason, I am sticking to the official school ages of 6-18 for the regressions. In addition to including age in my regressions, I am also including an age squared variable. As seen in Figure 8 and Figure 9, enrollment follows a fairly quadratic trend, although with very different means for each year. Because it follows this trend, it is necessary to include both age and age squared.

Sex and Location Variables:

'Female' is the dummy variable for sex, with one being equal to female and zero being equal to male. This is important to include because historically, there have been very different education trends for men and women in Liberia.

I am including the dummy variable 'urban' where one is people located in urban areas and 0 is people located in rural areas. The other location variables included in the dataset are counties, strata, and primary sampling units. There are 15 counties in Liberia. The 2007 dataset has 31 strata: an urban and rural area for each of the counties and a separate stratum for the capital Monrovia. 2019 has 30 strata, which are the urban and rural areas for each country without a separate stratum for Monrovia. The 2007 dataset has 300 primary sampling units, and the 2019 dataset has 325. I am including only the urban variable for location in the regressions because including any other location variables would cause perfect multicollinearity issues with other variables and the regression would not be able to be completed. The conflict data that I will be using sorts the conflict by county and urban/rural, which is the same as sorting by strata. The prewar education variable I will be using is constructed by giving a different value to each primary sampling unit. Because of this, I am not able to use country, strata, or primary sampling unit as a control. To measure household wealth, I am using the wealth index constructed by the DHS. The wealth index is a way to measure the standard of living. Each household is given a factor score depending on characteristics relating the household's socioeconomic status such as consumption and source of drinking water. This is a relative measure of wealth, with scores depending on the household's ranked depending on their factors compared to the other households. The scores are standardized using a normal distribution with a mean of 0 and a standard deviation of 1. The DHS dataset multiplied the standardized scores by 100,000 for data storage purposes, so I divided them by 100,000 to get the original standardized scores.

Education Loss:

In order to measure bounce-back, I need to figure out how much each individual should theoretically be able to recover from the war. I need to know how much each individual lost as a result of the war, so that I can know how much they would need to 'bounce-back' to the levels they would have had without the war. Since I am measuring enrollment, I need to know how much education people lost. The datasets do not include data on parent's levels of education, so for a measure of education loss, I am going to be using data from people who lived in the same area.

For 2007, I am going to use education trends from before the war to project how many years of education people would have had in a counterfactual world in which the civil wars never occurred. Because the trends for years of education were so different for men and women, I am a doing different projections for each sex. In 2007, people who were 36 years old would have been 18 in 1989 when the war first started, so theoretically their education levels should not have been

impacted by the war. So, I am going to use 36-51 year-olds in 2007 as the trend that I will project onto people who did lose education due to the war. I chose 51 year-olds as the cutoff because the trends for education shifted several times before the wars. 51 year-olds should have completed their education in 1974. I want to capture what was happening directly before the war.





Figure 10 show the average years of education for each age. I am using the war timeline on the x axis to illustrate the education trends before and after the wars. Year 0 on the war timeline is people who were 18 in 1989, so they would not have had their education impacted by the war. Everyone greater than year 0 had their education impacted by the war and people before year 0 did not. Year -20 of the war timeline is people who were 18 in 2003 when the war ended.

As seen by Figure 10, the years of education trend started changing a few years before the onset of the war (year 0 in the war timeline). This is likely because the economy began stagnating in the 1980s and the Doe government at the time was anti-elite, and therefore likely did not prioritize education. However, I cannot only include a few years before the war because the dataset would not include enough observations for each location to create a reliable trend with only a few years of data. In order to balance including the trend as close to the start of the war as possible and including enough observations to get a representative trend, I decided on 15 years of data, so 36-51 year-olds.

I am doing the trends by stratum. Theoretically, people in the same stratum should expect to get about the same years of education because they should be similar in other characteristics. I originally wanted to construct the trends by primary sampling unit. I feel that people in each primary sampling unit are going to be more similar than people in each stratum, because each primary sampling unit is much smaller. However, there were not enough observations for each primary sampling unit to get a representative trend, so I am using strata instead.

As mentioned before, the education trends were very different for men and women, so I am doing a separate trend for each of them.



Figure 11: Women's Years of Education Trend

The women's trend before the war was fairly linear and improving. The average years of education was about 3.5 before the beginning of the war. Because the trend was linear, I used a simple linear regression to predict how the trend would have continued in the counterfactual world.



Figure 12: Graph of Predicted vs Actual Average Years of Education by Age for Women

Figure 12 is the graph of my projection of years of education for 19-35 year-old women using data from 36-51 year-olds, compared to the actual years of education for 19-35 year-olds. The predicted years of education is generally higher than the actual.



Figure 13: Years of Education Trend for Men

The men's education trend was beginning to level off before the start of the war. The average years of education for men was almost double that of women before the start of the war. Because the trend was not linear, instead of a simple linear regression, I regressed log(age) onto years of education to predict the trend.


Figure 14: Graph of Predicted vs Actual Average Years of Education by Age for Men

Figure 14 is the predicted years of education for men compared to the actual years of education. Compared to the prediction for women, men were expected to have a much flatter education trend. This is because the years of education for men began to level off before the beginning of the war.

After predicting the counterfactual years of education, I subtracted the actual years of education from the predicted years in order to get the years of education lost due to the war. I could not use years of education for current students, because it is not known what their total years of education will be. It doesn't make sense to predict that a 6-year-old should have 6 years of education. So, I only used actual years of education for 18-35 year-olds.

For each stratum, I took the mean years of education loss and assigned that average to 6-18 year-olds as the education loss variable. This is effectively a measure of how much recovery in education years an individual could expect to see in the post-war period based on their stratum and gender. This the measure that I am using for bounce-back.

For males, there was an average of 7.78 years of predicted education. Meaning that on average, without the war we would have expected men to get 7.78 years of education. There was actually an average of 6.52 years of education for men, meaning that men are predicted to have lost a little over a year of education due to the war on average.

For females, there was an average of 4.61 years of predicted education. So, without the war, women theoretically would have received about 4.61 years of education. They actually received an average of 3.98 year of education. Women lost about 2/3 of a year of education.

A year of education is a lot, especially considering women had less than four years total. However, it is not as big of a loss as one might expect during a war that was so destructive. This is likely because the trends were beginning to level off before the war due to the political and economic turmoil. Had the trend been consistently linearly increasing, we would have expected more predicted years of education and therefore more education loss. It also makes sense that men would have lost more education than women, because they had more than double the amount of education as women before the war, so they had more to lose.

For 2019, I used the same method, except I changed the ages to be reflective of people's ages in 2019 rather than 2007. For the trend to predict years of education. I am using 48-63 year-olds. 48 year-olds in 2019 were 18 in 1989. I am using the actual years of education for 31-45 year-olds, so these should be the same age cohorts as the people in 2007.

For males, there was an average of 8.59 years of predicted education and 7.92 years of actual education. So, there was a little more than half a year of education loss for males due to the war. This is a higher amount of predicted and actual years of education and less education loss than we found in 2007, even though it should be the same group of people. What I expect is happening is that education is correlated with health and life expectancy. The people we are using to predict the trend in 2019 are 48-63 year-olds, it is likely that people who are more highly educated are more likely to live to the age included in the sample, and therefore educated men are overrepresented.

For females, there was an average predicted 3.85 years of education and an actual average of 4.38 years of education. So, the education loss is actually negative, implying that women gained about half a year of education that they would not have gotten in the counterfactual world without the war. This may be due to a change in preferences over time. It may be that wealthier families did not educate their daughters before the 1980s, but over time the preferences changed, and they began educating their daughters, so the trend we used to predict did not take the change in preferences into account.

Prewar education:

In order to use the measure of bounce-back I just laid out in education loss, I also need to control for prewar levels of education in each area. If I do not control for prewar levels of education, in my education loss variable I will be capturing both education loss (which is our bounce-back measure) and the general level of education for each area. The general level of education in area may have effects on enrollment outside of bounce-back. For example, it is possible that areas with high levels of education before the war expected their children to be educated and prioritized reopening schools more than areas with lower levels of education. In order to properly measure bounce-back, I need to control for those effects and create a measure of education for each area before the war.

Similarly to the education loss variable, I decided instead to use the education information for the people in each area who were old enough to not have their education impacted by the war, to estimate the prewar education levels for each 6-18-year-old. For areas I used primary sampling units. Because I am only using an average rather than a getting a trend, I do not need as many observations as I did for the education loss variable, so it was practical to use primary sampling units for this variable. I also estimated education levels separately for men and women, because of significant differences in the education levels and trends before the beginning of the war as mentioned before (Figure 11, Figure 13).

For 2007, the were 298 primary sampling units (psu). I took the mean of the years of education for 36-51 year-olds in each psu by gender and assigned that value as the prewar education level for 6-18 year-olds. There were 43 psus for women and 33 psus for men that had less than 5 observations. I decided that these were not large enough samples to get a representative estimate for the level of education, so for the psus with less than five observations, I used everyone in the stratum instead of in the psu. There were 31 strata in the survey. There was an urban and rural stratum for each of the 15 counties as well as a stratum for the Greater Monrovia area (the capital of Liberia).

There were 4,801 36-51 year-olds in the sample. As before, I decided to use 36-51 yearolds because people who were 36 in 2007 were 18 in 1989 when the war started. This means everyone 36 and above should not have had their education significantly impacted by the wars. I did not want to use everyone above 36-years-old, because as you go further out the trends begin to change and there is much more variation due to the smaller sample size of older people. 15 years was enough to get a representative sample for most psus, but not so much that the data started getting less clean because of variation.

Figure 15 is the graph of years of education for women by age. The ages highlighted in red are the 36-51 year-olds used in this measure.



Figure 15: Years of Education for Women by Age in 2007

Figure 16 is the same graph, but for men.



Figure 16: Years of Education for Men by Age in 2007

The differences in Figures 15 and 16 demonstrate why I estimated the levels separately for men and women. The trends are significantly different, with women have a steadily increasing linear trend and men having a logistic trend that is beginning to level off just before the war.

For males, there was an average of 6.94 years of prewar education. This means on average, men had 6.94 years of education right before the war began. For females, there was an average of 2.75 years of prewar education. This means that on average, women had 2.75 years of education right before the war began, which is less than half of what the men had.

For 2019, there were 325 psus. Just like in 2007, I took the mean of years of education for 48-68 year-olds in each psu by gender and assigned that value to the 6-18 year-olds in the same psu. There were 60 psus for women and 70 psus for men that had less than 5 observations, so like before, I replaced those with data from the stratum rather than the psu. There are 30 strata in this dataset, an urban and rural area for each of the 15 counties, without a separate stratum for Monrovia.

There were 4,820 observations for 48-68 year-olds. 48 year-olds in 2007 were 18 in 1989, so everyone older than 48 should not have had their education greatly impacted by the war. I used 20 years instead of 15 in this case because there were not enough observations for each psu if I used less than 20.

Figure 17 is the graph of years of education for women by age. The ages highlighted in red are the 48-68 year-olds used in this measure.



Figure 17: Years of Education for Women by Age in 2019

Figure 18 is the same graph for men.



Figure 18: Years of Education for Men by Age in 2019

There is more variation than is ideal for this age group because there are less observations as age increases; however, I do not believe the variation is extreme enough for it to cause an issue.

Using the 2019 data, men had an average of 6.99 years of prewar education, and women had an average of 2.55 years. These are fairly similar numbers to the averages I found using 2007 data (6.94 years and 2.75 years respectively). The numbers for women are slightly more different, but not so far off that it raises concerns.

Predicted Education Gain:

I am also including a variable measuring how much we expected each person to gain in years of education in a counter-factual world in which the war never occurred. Without including this variable, the education loss variable will be picking up the effect of both the loss in education due to the war and the amount of education people were projected to gain without the war. Controlling for predicted education gains allows the education variable to capture only the effect of the years of education lost due to the war.

To construct the predicted gains in education, I will be using the same years of education projection that I used to find the education loss. The projected years of education in a counterfactual world without the war was found using years of education for 36-51 year-olds in 2007 and 48-63 year-olds in 2019 regressed onto age for females and log(age) for males separately for each strata. For the education gains variable, I took this predicted years of education value for each 19-35 year-old and subtracted their prewar education variable. The prewar education variable was constructed by gender and primary sampling unit.

The predicted education gains variable is constructed the same way as the education loss variable, except I am subtracting years of prewar education rather than actual years of education for each individual. Because of this, there is a fairly high correlation between the predicted education gains and the education loss variables. The correlation between these variables for 6-18 year-olds in 2007 was .899. In 2019, the correlation was .927. These correlations do not make the estimates biased, but they may make it difficult for the model to differentiate between the effects of the two variables.

For females in 2007, the mean predicted gain in years of education was 1.88 years. The mean for males was 0.17 years. This discrepancy makes sense because the male education trend

was leveling off before the war more than the female trend, so the male trend was not predicted to greatly increase even in the counterfactual world without the wars. The female trend was still increasing at a fairly linear trend, so it makes sense that females would have more predicted years of education gain. The variable had fairly high variation of variation, with a standard deviation of 2.03 for females and 2.63 for males.

In 2019, the mean predicted gain in years of education for females was 1.64 years, instead of the 1.88 years in 2007. For males, the mean was 0.28 years in 2019 compared to the 0.17 years in 2007. The standard deviation for females was 1.86 compared to 2.03 in 2007. The standard deviation for males in 2019 was 3.22 and 2.63 in 2007. These means and standard deviations are not extremely different, which indicates that the samples in 2007 and 2019 were not significantly different in terms of predicted education gain.

Conflict:

The dataset I used to measure conflicts and death was the Georeferenced Event Dataset (GED) from the Uppsala Conflict Data Program (UCDP). The GED is a dataset of all conflicts from 1975-present that were reported. There are 548 conflicts in Liberia between 1989 and 2003 in the dataset. It gives the approximate coordinates of the conflict sites, so I used the coordinates to sort each event into the county it occurred in and whether it was an urban or rural area. For urban areas, I used the capital of each county and the large cities mentioned on the county pages of Liberia's Ministry of Internal Affairs website⁴⁸. I treated everywhere else as rural.

I wanted do conflict by primary sampling unit, as I think that would be a better representation of the effect that conflict had on the immediate surrounding community; however,

⁴⁸ ("Overview of Liberia," n.d.)

the DHS dataset does not give location data for their primary sampling units due to privacy concerns, so conflict by strata was the most specific I could do.

I am trying to use level of conflict as a measure of destruction, both in physical and human capital in a given area. Areas with more destruction have to recover more than areas little destruction. In 2007, I would expect areas that were more greatly devastated would have lower enrollment, as they would take longer to rebuild and get kids back into school.

There were 22,634 estimated deaths recorded in the dataset. An estimated 200,000 people died in the civil wars, so this dataset does not record all of the deaths or events that actually occurred. This is because not all of the conflicts were recorded. It is not possible to get data for every single conflict that happened within the war, but this is the most extensive dataset. There is no reason to believe that the data is biased towards a certain area, so I am treating this data as a representative sample of the conflict and deaths, but with the caveat in mind that it is a sample rather than an exhaustive list.

I originally decided to split the conflict data into two periods: 1989-1996 and 1997-2003. As seen by Figure 19, there was a big drop-off in conflict between 1997 and 2000.

Figure 19: Conflict Event by Year



This drop-off approximately coincides with the end of the official first civil war and the beginning of the second. I originally decided to split the data into two periods to see if they differently affected enrollment. I would expect that if an area had a lot of conflict in the early period and very little in the later one, they would have had time to somewhat rebuild the area earlier than other areas were able to, so they would not be as affected by the conflict by 2007.

I also expected the proportion of conflicts or 'events' that each area had to be approximately the same as the proportion of deaths, but that was not the case everywhere.



Figure 20: Death, Conflict Event, and Population by County

Lofa County is a good example of this. It has only 10.04% of the population, but 17.95% of the conflict events and 22.41% of the deaths.

Because of the disparity between deaths and conflict events in some places, I also decided to include both deaths and events to see which had a greater impact. So, I originally had four conflict variables: early deaths (1989-1996), late deaths (1997-2003), early events, and late events.

However, after looking at the correlations with the conflict variables, it seems like I would have a near-perfect multicollinearity problem if I included all four variables.

Figure 21: Correlation Between Deaths and Events

. corr deaths events
(obs=34,670)

	deaths	events
deaths	1.0000	
events	0.9587	1.0000

While there was some variation in deaths and conflict events, I do not think it is enough variation for a model to include both and be effective. There is a correlation of .9587 between events and deaths, which is likely too much for the model to effectively differentiate between them.

	earlyd~s	latede~s	earlye~s	lateev~s
earlydeaths	1.0000			
latedeaths	0.8616	1.0000		
earlyevents	0.9226	0.8978	1.0000	
lateevents	0.8550	0.9854	0.9054	1.0000

Figure 22: Correlation Table for Early Deaths, Late Deaths, Early Event, and Late Events

There is more variation in deaths and events in the earlier period (a correlation of .9226) than in the later period (a correlation of .9854), but I still think even the early period does not have enough variation.

I believe that deaths should ultimately be more impactful than events, because I am trying to capture destruction. Deaths definitely captures loss in human capital more than conflict events, as events had anywhere from 0 to 1,647 deaths. A conflict with 1,647 deaths will have a much greater negative impact on the community than a conflict with 0 deaths.

There is a correlation of .8616 for deaths in the early period and deaths in the later period. While this is more variation than we have seen before, it may not be enough to avoid nearperfect multicollinearity problems, so it does not seem useful to include two different time periods.

While collinearity does not make an estimate biased, it can make it impossible to differentiate between which variables are having an effect on the dependent variable. If I included multiple conflict variables, it would be difficult to tell which is having an effect on enrollment; therefore, I am only using deaths as a measure of conflict. That way it will be easy to tell what effect the conflict is having on enrollment without trying to differentiate between multiple conflict variables. Because of this near-perfect multicollinearity issue, we will not know whether the variable deaths is truly capturing only the effects of deaths, or also capturing the effects of conflict events as well. We will also not be able to tell whether deaths or events in the early period or the later period are driving the effects of the death variable.

Aid:

The most important variable I am omitting is a measurement of foreign aid in Liberia after the civil wars. As I alluded to in the background section, it was impossible to keep track of the amount of aid flowing into Liberia in the rebuilding period. Not only were most major governments and NGOs giving aid, but so were countless small organizations that were working with smaller budgets and focusing on smaller areas.

I was originally planning on using aid as one of my main explanatory variables; however, I was not able to find a good way to estimate the amount of aid. Even the larger donors like USAID and the UN only have collective data at a country level, rather than organized information about specific projects in specific places. Some organizations post information about specific projects, but I would have to go through each aid organization's postings for each project in Liberia to create a master dataset that includes funding and projects for each location within Liberia, which is not feasible in the scope of this project. Country level data is not helpful to my analysis, as there is no way to differentiate in the data who benefitted from what aid.

Because of the unavailability of aid data, I have not been able to include it and control for it in my analysis, which means that other variables are likely picking up the effects of aid. I anticipate that aid will be highly correlated with deaths, as theoretically aid groups will likely be targeting the areas most devastated by war. Because of this correlation, it is likely that deaths will be picking up both the effects of destruction and, particularly in 2019, the effects of aid. I also anticipate that female will be somewhat correlated with aid, as a lot of the aid program were aimed at increasing gender equality in education. It is likely most enrollment change over time for females is due to those programs. Lastly, it is possible that urban is also correlated with aid. For the major donors, it is much easier to set up programs in urban areas than rural areas, as you are reaching more people. It is possible that urban areas received higher levels of aid because of this. It is also possible, however, that smaller donors and groups filled the need in some of the rural areas, so aid is not highly correlated with urban.

Bounce-back:

I am attempting to measure bounce-back. On a macro-level, complete bounce-back would occur if enrollment levels returned to either the enrollment rates before the war, or the enrollment rates predicted in a counterfactual world without the war, depending on how you want to define it. Complete bounce-back would occur if enrollment rates returned to normal levels not only on a national level, but for each group. For example, imagine a scenario in which enrollment rates returned to prewar levels overall, but places with a lot of destruction did not recover to prewar levels. Areas with low levels of destruction had disproportionate recovery and their levels are much higher than they were before the war. These areas' enrollment rates even out so that the overall national level is the same as it was before the war. This is not complete bounce-back because the high-destruction areas did not fully recover. People, groups, or areas with more loss or destruction have to recover proportionately to people with less loss for complete bounce-back to occur.

In order to measure this macro-level bounce-back in Liberia, I would need enrollment levels from before the war, as well as for a few years before the war to get an idea of the trend. I would also need data from during the war and directly after the war in 2003 rather than in 2007, because by 2007 some bounce-back will have already occurred, so it is hard to precisely measure recovery. I would also ideally have census data with enrollment rates rather than samples in order to get a good idea of recovery with more precise locations than just at the strata level. Because this data does not exist, I cannot accurately measure bounce-back on this overall macro-level, but I can measure it on a micro-level. By this, I mean I cannot measure whether everyone recovers to their prewar levels, but I can measure whether or not people recovered proportionately to others.

Education loss is a good indication of this micro bounce-back. When controlling for all of the other explanatory variables, education loss measures how many years of education each area is predicted to have lost during the war. That education loss measure indicates how much each person needed to recover to get back to what their predicted levels of education were in the counterfactual world without the wars.

If the coefficient for education loss is positive, and an increase in years of education loss is associated with an increase in the probability of being enrolled, it would imply that areas with more loss experienced disproportionately more recovery than areas with less loss. This would be possible if recovery programs specifically targeted high-loss areas and were very effective. In this scenario there would be more than complete bounce-back (on this micro-level) because not only did areas with more loss recover at proportional rates to areas with less loss, but they actually recovered at disproportionately high rates. If the education loss coefficient is negative, and an increase in years of education loss is associated with a decrease in the probability of being enrolled, it would imply that areas with more loss did not recover as much relative to areas with less loss. This is the most intuitive outcome, because it is likely that areas with more loss would struggle to recover more than area with less loss would struggle to recover. It would imply that significant bounce-back did not occur, because people with more loss did not recover proportionately to people with less loss. If the effect of education loss was 0, this would imply micro-level bounce-back. Controlling for other variables like initial levels of education before the war and how much each area was expected to gain without the war, if there was no effect for education loss, then it would not matter how much a person lost due to the war. This would imply that people recovered proportionately regardless of their levels of loss, and this would therefore indicate micro-level bounce-back.

Change over time:

Variable	Mean	Standard Deviation	Min	Max	Observations
enrolled	0.510	0.500	0	1	11,466
female	0.497	0.500	0	1	11,466
age	11	3.6	6	18	11,466
agesq	134	84.5	36	324	11,466
urban	0.411	0.492	0	1	11,466
wealthindex	0.269	1.08	-1.31	3.98	11,466
deaths	2232	2119	0	5099	11,466
prewar_educ	5.12	3.55	0	14	11,466
predgain	1.00	2.05	-8.87	10.59	11,466
educloss	0.94	1.99	-7.68	11.43	11,466

Table 1: Table of Means for 2007 Ages 6-18

Table 1 is the table of means for the variables I will be including in my analysis for 2007. Notably, about 51% of 6-18 year-olds were enrolled in 2007, and about 41% of 6-18 year-olds lived in urban areas. The standard deviation for deaths and the education variables are fairly high. This suggests that there was a lot of variation in the number of deaths by area, which is consistent with the conflict data that I discussed previously. It also suggests that there was a good amount of variation in the education variables. This makes senses because all three education variables were manually created using estimates or predictions by location. Some of the locations do not have many observations, so there is a fair amount of variation between locations.

Variable	Mean	Standard Deviation	Min	Max	Observations
enrolled	0.810	0.392	0	1	14,633
female	0.497	0.500	0	1	14,633
age	11	3.6	6	18	14,633
agesq	142	84.1	36	324	14,633
urban	0.584	0.493	0	1	14,633
wealthindex	0.497	1.13	-2.51	4.14	14,633
deaths	2072	2154	0	5099	14,633
prewar_educ	4.84	3.54	0	16.14	14,633
predgain	1.25	2.55	-9.84	6.84	14,633
educloss	0.03	2.36	-11.33	6.00	14,633

Table 2: Table of Means for 2019 Ages 6-18

The biggest difference between 2007 and 2019 is the overall enrollment rates. 81% of kids 6-18 are enrolled in 2019 vs. 51% in 2007. There are also more kids living in urban areas in 2019 (58.42%) than in 2007 (41.09%). It seems like a lot of kids moved into urban areas after the war, possibly due to the fact that a lot of kids were orphaned. There is also an education loss of .029 years on average in 2019 vs an average of .943 years in 2007. As discussed when constructing the education loss variable, it is possible that this is due an overrepresentation of some groups based on who is still alive from the older age cohort. There are also slightly greater standard deviations for the education variables in 2019. This is likely due to the fact that there

were less observations in 2019 than in 2007 for the age groups I used to create the education variables.

I am examining estimates of how enrollment rates changed over time for various groups. Some of the estimates of the mean of enrollment are missing standard errors. This is because in order for Stata's survey command to estimate standard errors there needs to be multiple primary sampling units in each stratum. All of the strata have more than one primary sampling unit, but not all of primary sampling units have observations for every group. For example, several of the primary sampling units do not have observations for every quintile of wealth index. Because of this, some of the strata are left with only one primary sampling unit with observations when I am estimating the mean of enrollment for a certain quintile. Because of this, the survey command is unable to estimate standard errors because there is no variation within a stratum.

The differences between the groups or years are found using a regression of that group/year's dummy variable on the enrolled variable.

Enrollment in 2007	Girls	Boys	Difference
Mean	0.482	0.538	0.056
Standard Error	0.0168	0.0150	0.0137
Number of Observations	5,672	5,794	11,466

Table 3: Mean Enrollment for Girls vs Boys in 2007

In 2007, 48.21% of females aged 6-18 were enrolled in school compared to 53.78% of males aged 6-18. This is more than a 5-percentage-point difference and is statistically significant at the 1% significance level. This means that there are significantly more boys that are enrolled in school directly after the war then girls.

Enrollment in 2019	Girls	Boys	Difference
Mean	0.814	0.807	0.007
Standard Error	0.0106	0.011	0.0087
Number of Observations	7,008	7,625	14,633

Table 4: Mean Enrollment Girls vs Boys in 2019

By contrast, in 2019 81.40% of females are enrolled and 80.69% of males are enrolled. There are actually more estimated females enrolled than males, although there is not a significant difference in the enrollment levels for girls and boys in 2019. It seems that in the years 2007-2019, Liberia closed the enrollment gap between girls and boys. Considering there was a 5 percentage-point difference in 2007, it is impressive to close that gap in just 12 years. This is likely due to the fact that a lot of the education aid program were targeted towards promoting gender equality in education, so the girls in Liberia likely saw a larger benefit from the aid in the rebuilding era.

Enrollment for Girls	2007	2019	Difference
Mean	0.482	0.814	0.332
Standard Error	0.0168	0.011	0.0198
Number of Observations	5,672	7,008	12,680

Table 5: Mean Enrollment for Girls in 2007 vs 2019:

Enrollment rates for females improved from 48.21% in 2007 to 81.40% in 2019. That is a 33.19 percentage-point increase over 12 years, and it is statistically significant at a 1% level.

33.19 percentage-points is a huge increase and indicates that the rebuilding of education systems in Liberia were extremely successful.

2007	2019	Difference
0.538	0.807	0.269
0.0150	0.011	0.0188
5,794	7,625	13,419
	0.538 0.0150 5,794	2019 0.538 0.807 0.0150 0.011 5,794 7,625

Table 6: Mean Enrollment for Boys in 2007 vs. 2019

Enrollment rates for males increased from 53.78% in 2007 to 80.69% in 2019. That is a 26.91 percentage-point increase, and it is statistically significant at a 1% level. While it is not as large of an increase as the females, it is still an incredibly large increase in enrollment rates.

Table 7: Enrollment for Rural vs Urban Areas in 2007

Enrollment in 2007	Rural	Urban	Difference
Mean	0.407	0.658	0.250
Standard Error	0.0227	-	0.0263
Number of Observations	6,430	5,036	11,466

In 2007, 40.73% kids living in rural areas were enrolled, while 65.75% of kids in urban areas were enrolled. This is a 25.02 percentage-point difference between urban and rural areas, and it is a statistically significant difference. This makes sense, as urban areas were likely able to recover more quickly than rural areas as they had more resources and were likely prioritized in terms of rebuilding, as the majority of the economic activity occurs in urban areas.

Enrollment in 2019	Rural	Urban	Difference
Mean	0.739	0.861	0.123
Standard Error	0.0155	0.014	0.0208
Number of Observations	8,948	5,685	14,633

Table 8: Enrollment for Rural vs Urban Areas in 2019

In 2019, 73.89% of kids in rural areas were enrolled in school, whereas 86.14% of kids in urban areas were enrolled in school. This is a 12.27 percentage-point difference between urban and rural areas, and it is a statistically significant difference. While still a significantly large difference, it is down from a 25.02 percentage-point difference in 2007. This means Liberia has effectively halved the difference in enrollment between urban and rural areas in 12 years. This is incredibly impressive and implies that the aid and rebuilding programs were not only targeted towards urban areas. It implies that they were targeted towards both, but actually towards rural areas more often than urban areas, which is the opposite of what I expected, as I expected it would be easier for donors to target urban areas.

Enrollment in Rural Areas	2007	2019	Difference
Mean	0.407	0.739	0.331
Standard Error	0.0227	0.0155	0.0275
Number of Observations	6,430	8,948	15,378

Table 9: Enrollment in Rural Areas in 2007 vs 2019

Enrollment rates in rural areas increased from 40.73% in 2007 to 73.87% in 2019. That is a 33.14 percentage-point increase and is statistically significant at the 1% level. This indicates that rebuilding programs that encouraged enrollment in rural areas were successful.

Enrollment in Urban Areas	2007	2019	Difference
Mean	0.658	0.861	0.204
Standard Error	-	0.0138	-
Number of Observations	5,036	5,685	10,721

Table 10: Enrollment in Urban Areas 2007 vs 2019

Enrollment rates in urban areas increased from 65.75% in 2007 to 86.14% in 2019. That is a 20.39 percentage point increase. We do not know if this statistically significant because of the primary sampling unit issue mentioned before, but it likely is because of how large the difference is. While still a huge increase, it is a smaller percentage point increase than rural areas experienced (33.14 percentage-points). This could indicate two things. The first is that rebuilding programs were targeted towards rural areas more than urban areas, which allowed rural areas to experience more growth in enrollment rates. The second is that urban areas started at a much higher rate already, and they could be approaching their ceiling. A 33.14 percentage point increase were enrolled in school in 2019, which is unrealistic to achieve. It would not have been possible for urban areas to improve much more rural areas. Because rural areas started at much lower rates, they were able to improve more.

I also want to see how enrollment rates changed over time for the different levels of wealth. The wealth index from the DHS splits households into quintiles for me, so I am going to use each quintile for each year.

Enrollment by Wealth in 2007	Mean	Number of Observations
Poorest Quintile	0.302	2,206
Poorer Quintile	0.385	2,118
Middle Quintile	0.495	2,300
Richer Quintile	0.539	2,412
Richest Quintile	0.744	2,430

Table 11: Enrollment Rates by Wealth Quintile in 2007

In 2007, level of enrollment decreases with every wealth quintile. 74.44% of the richest quintile are enrolled, then it is 53.86%, 49.46%, 38.53%, and 30.20% respectively. That is a 44.24 percentage-point difference in the top quintile compared to the bottom quintile. It is clear that directly after the war, wealthier people were able to get their kids in school faster than less wealthy people. This may be because wealthier people could afford to send their kids to schools that were further away or more expensive, and since there was a limited supply of schools, wealthier kids were overrepresented. It could also be because poorer households needed their children to stay home and work or help the household because they were so negatively impacted by the war.

Enrollment by Wealth in 2019	Mean	Number of Observations
Poorest Quintile	0.662	3,848
Poorer Quintile	0.759	3,799
Middle Quintile	0.804	3,313
Richer Quintile	0.885	2,099
Richest Quintile	0.921	1,574

Table 12: Enrollment Rates by Wealth Quintile in 2019

In 2019, all of the quintiles had significant gains in enrollment compared to 2007. In the top quintile 92.05% of kids are enrolled in 2019, which is a 17.61 percentage-point increase. In the second-wealthiest quintile 88.45% of kids are in enrolled, which is a 34.6 percentage-point increase from 2007. For the middle quintile, 80.43% of kids are enrolled, which is a 30.97 percentage-point increase. In the second-poorest quintile, 75.94% of kids are enrolled, which is a 37.41 percentage-point increase from 2007. Finally, 66.22% of kids in the bottom quintile are enrolled, which is a 36.02 percentage-point increase.

The percentage-point increases are fairly consistent across quintiles except for the top quintile, in which it is much less. It makes since that the increase is far less in the top quintile because the enrollment rates already started very high, so there was not as much room for improvement. Because enrollment rates started so much lower in 2007 in the lower quintiles, had the rebuilding programs equally targeted everyone, the lower quintiles likely should have seen more improvement than the higher quintiles, because they were able to see more improvement. This did not happen. The lower quintiles did not close in on the upper quintiles in terms of enrollment rates, they just all improved at around the same levels. This implies that rebuilding programs did not target lower income people and may have actually benefitted wealthier people more.

I am comparing areas that experienced high levels of conflict to areas that experienced less conflict. I am defining high conflict areas as areas that experienced over 1000 deaths. Of the 31 strata, 8 of them experienced over 1000 deaths.

Enrollment in 2007	Low-Conflict Area	High-Conflict Area	Difference
Mean	0.452	0.558	0.106
Standard Error	0.0257	0.0154	0.0230
Number of Observations	6,709	4,757	11,466

Table 13: Low-conflict vs High-conflict Enrollment in 2007

In 2007, 45.17% of kids in low-conflict areas were enrolled, while 55.76% of kids in high-conflict areas were enrolled. This is a 10.59 percentage point difference and is statistically significant. This is the opposite of what I would expect to be true. I would assume that areas which experienced the highest degree of conflict would have lower enrollment rates soon after the war because there would have been more devastation in the area, and it would have been more difficult to get back to school.

There are a few reasons this could have happened. The first is that areas that experienced more conflict are also correlated with something else that raises enrollment rates, like wealth. It is possible that this is the case and areas with high-conflict did experience more destruction and loss more enrollment, but they started at a significantly higher level of enrollment than low-conflict areas and the difference would be significantly bigger without the war, so we cannot see the effects by only looking at 2007. Another possibility is that in the four years in between the end of the war and the DHS survey, a disproportionate amount of resources were poured into the

communities with the most devastation, so they had some recovery before the beginning of the survey, and their enrollment levels would have been lower than the low-conflict areas in 2003 or 2004 directly at the end of the war.

Enrollment in 2019	Low-Conflict Area	High-Conflict Area	Difference
Mean	0.804	0.817	0.013
Standard Error	0.0111	0.0173	0.0205
Number of Observations	10,219	4,414	14,633
Number of Observations	10,219	4,414	1

Table 14: Low-conflict vs High-conflict Enrollment in 2019

In 2019, 80.44% of kids in low-conflict areas are enrolled and 81.70% of kids in highconflict areas are enrolled. There is not a statistically significant difference between the enrollment rates for high-conflict and low-conflict areas in 2019. This could mean that aid and rebuilding resources were not targeted towards areas that experienced the most conflict like I assumed.

2007	2019	Difference
0.452	0.804	0.353
0.0257	0.0111	0.0280
6,709	10,219	16,928
	2007 0.452 0.0257 6,709	200720190.4520.8040.02570.01116,70910,219

Table 15: Enrollment in Low-conflict Areas in 2007 vs 2019

Enrollment in low-conflict areas increased from 45.17% in 2007 to 80.44% in 2019, which is a 35.27 percentage-point increase and it is statistically significant at the 1% level.

Enrollment in High-Conflict Areas	2007	2019	Difference
Mean	0.558	0.817	0.259
Standard Error	0.0154	0.0173	0.0231
Number of Observations	4,757	4,414	9,171

Table 16: Enrollment in High-conflict Areas in 2019

Enrollment increased in high-conflict areas from 55.76% in 2007 to 81.70% in 2019. This is a 25.94 percentage-point increase, and it is statistically significant. It is less than the percentage-point increase experienced in low-conflict areas. This is not what was expected to happen. It implies that the rebuilding efforts had a bigger impact on areas of low-conflict than areas of high-conflict.

Methods

The main methods of analysis I will be using are probit models and linear probability models.

Linear Probability Models:

Linear probability models are standard OLS regression models with a dummy variable as the dependent variable. Their results are generally easier to interpret than the probit models.

The basic equation for the linear probability model is:

$$P(y = 1|x) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon$$

Where x is the explanatory variables and E(y|x) is the probability of 'success', which in this case is the probability of being enrolled in school. $\hat{\beta}_0$ is the probability of success if all of the independent variables are equal to 0. $\hat{\beta}_i$ is the slope coefficient and it measures the predicted change in the probability of success with a one unit increase in x_i . Holding all other explanatory variables constant, $\Delta P(y = 1|x) = \beta_i \Delta x_i$.

However, there are some issues with using a linear regression. One of the main issues with a linear probability model is the fact that it is linear. This means that you can get probability estimates that are less than 0% or greater than 100%. Linear probability models are generally fairly accurate in the middle of the regression line when the values of independent variables are close to the mean, but as you approach 0 or 1, the estimates get worse.

Linear probability models are also by nature heteroskedastic. This is because the dependent variable can only take on the values of 0 or 1, but the predicted values will be anywhere from 0 to 1, so the value of the residuals change with changing explanatory variables. The variance for a linear probability model is Var(y|x) = P(y = 1|x)[1 - P(y = 1|x)].

OLS regression models are based on the assumption of homoskedasticity, which is that assumption that the variance of the unobserved error, u, conditional on the explanatory variables X, is constant.

$$Var(u|x_1, x_2, \dots, x_k) = \sigma^2$$

When the homoskedasticity assumption does not hold, the variances are biased. A heteroskedastic model is not the best linear unbiased estimator like an OLS model is.⁴⁹ The linear probability model is always heteroskedastic.⁵⁰ So, the linear probability model is never the best

⁴⁹ (Wooldridge 2012)

⁵⁰ (Pedace 2016)

linear unbiased estimator. The coefficient estimates are still unbiased, but the standard errors are biased.

An effective way of correcting for heteroskedasticity in the linear probability model is to use robust standard errors, which I will be using in my linear probability models. I am using the survey command in Stata to add weights to my variables, and the survey command automatically applies robust standard errors.

Probit Models:

The linear probability model should give us a good estimate; however, because of the issues with the linear probability model, it is also necessary to use another model that corrects for some of these errors and compare the results. I will be using the probit model for this purpose.



Figure 23: Probit and Logit Models

The probit model is not linear, as seen in Figure 23 comparing the probit model to a logit model.⁵¹ Because of the non-linearity, it cannot give a probability estimate that is not between 0% and 100%. It is also more accurate on the fringes of the model, as Y approaches 0 or 1.

In order to create the probit model, we begin with the equation

$$P(y = 1|x) = G(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon)$$

Where G is a function that takes on values between 0 and 1. We will essentially using the linear probability model from before, but putting our observables through a function G. Probit models use the cumulative distribution function as G. The standard normal cumulative distribution function is just the area under the curve of the probability distribution function. The cumulative distribution function is represented through the equation:

$$G(z) = \phi(z) = \int_{-\infty}^{z} \phi(v) dv = \int_{-\infty}^{z} 2\pi^{-1/2} \exp\left(-\frac{v^{2}}{2}\right)$$

For each observation, our outcome (Y) is either 0 or 1. We are trying to predict the probability of success, but each observation does not actually tell us their probability it only tells us whether or not they were successful.

However, we assume there is a latent index, or an index that you cannot see, that all of the observations sit on. Where each observation lies on this index depends on how likely the person was to get an outcome of 1 or a success. At some point on this index is a crossing over point, where people get an outcome of 1. For some people, they were well above the threshold for

⁵¹ (Kumar 2023)

getting an outcome of 1 and for others they just barely passed the threshold. However, both of these people would have gotten outcomes of 1. The latent index can be represented by the equation:

$$y^* = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon, \ y = 1[y^* > 0].$$

If the latent index, or y^* , is greater than 0, then y (the probability of success) is equal to 1, and if y^* is less than 0 then y=0.

This means we do not know the error term for each individual observation because the error term is their outcome (0 or 1) minus their value on the latent index (which we do not know). We assume that these error terms are normally distributed.

Probit models are a Maximum Likelihood Estimator in which the model estimates the combination of β s that give the greatest probability of drawing your observed data. The maximum likelihood function is this:

$$f(y|x_i;\beta) = [\phi(X'\beta)]^{y} [1 - \phi(X'\beta)]^{1-y}.$$

Taking the log of the maximum likelihood function gives us the log-likelihood function:

$$\ell_i(\beta) = \log(f(y|x_i;\beta)) = y_i \log(\phi(X'\beta)) + (1-y_i)\log(1-\phi(X'\beta))$$

The log-likelihood sums the log-likelihood function for all observations:

$$\mathcal{L}(\beta) = \sum_{i=1}^{n} \ell_i(\beta)$$

The maximum likelihood estimator is the combination of β s that maximizes the log-likelihood.

The probit model I will be using is this:

$$\begin{aligned} Prob(enrolled &= 1|X)_i \\ &= \phi(\beta_0 + PC_1x_1 + wealth_2x_2 + deaths_3x_3 + educationloss_4x_4 \\ &+ predictedgain_5x_5 + prewareducation_6x_6 + \mu) \end{aligned}$$

Where ϕ is the standard normal cumulative distribution function and PC is personal characteristics: age, age squared, female, and urban.

The results of probit models are somewhat difficult to interpret. The coefficient of β does not give the marginal effect of an independent variable like linear probability models, probit models give the change in the z score. The change in the probability estimate based on the same change in z-score depends on where an individual lies on the cumulative distribution function. The same change in z-score will be a greater change in probability of Y for a person near the center of the cumulative distribution function as compared to a person near the edge of the function. To make the interpretation of the results easier I will be using the average marginal effects. This calculates the marginal effects for each observation and averages them.⁵².

Results

All of the following models are using kids aged 6-18.

⁵² (Wooldridge 2012)

lumber of stra	ata = 30			Numb	er of obs =	11,466
umber of PSUs	= 298			Popu	lation size =	1,134,638
			Desi	gn df =	268	
				F(6,	263) =	121.14
				Prob	> F =	0.0000
		Linearized				
enrolled	Coefficient	std. err.	t	P> t	[95% conf.	interval]
female	2440683	.038246	-6.38	0.000	3193692	1687674
age	.7673896	.0473794	16.20	0.000	.6741065	.8606727
agesq	0255652	.0018776	-13.62	0.000	0292619	0218686
urban	.1820649	.0818089	2.23	0.027	.0209951	.3431348
wealthindex	.3544911	.0381161	9.30	0.000	.2794461	.4295362
deaths	-9.01e-06	.0000205	-0.44	0.660	0000493	.0000313
	and an end of the second					

Figure 24: Probit Model without Education Variables 2007

Figure 25: Average Marginal Effects

Average marginal effects

Number	of	strata	=	30
Number	of	PSUs	=	298
Model \	/CE:	Linea	riz	zed

Number of obs	=	11,466
Population size	=	1,134,638
Design df	=	268

Expression: Pr(enrolled), predict()
dy/dx wrt: female age agesq urban wealthindex deaths

	dy/dx	Delta-method std. err.	t	P> t	[95% conf.	interval]
female	0715595	.0110894	-6.45	0.000	0933929	0497262
age	.2249945	.0111592	20.16	0.000	.2030238	.2469653
agesq	0074956	.0004716	-15.89	0.000	0084241	0065671
urban	.0533805	.0238091	2.24	0.026	.0065039	.1002571
wealthindex	.1039349	.0108569	9.57	0.000	.0825592	.1253106
deaths	-2.64e-06	6.02e-06	-0.44	0.661	0000145	9.21e-06
Before running the full model including prewar education, predicted education gain, and education loss, I am running a model without those three variables to get an idea of the significance of the female and urban variables. Prewar education, predicted education gain, and education loss were constructed using data by gender and location. Because of that, a lot of the effect of female and urban are being picked up in those variables. For example, one of the major reasons the female variable would be negative is if there were attitudes discouraging girls from going to school. However, prewar education is education levels before war, and it would pick up that same attitude.

Without the education variables, female and urban are both significant. Being female is associated with a 7.16 percentage-point decrease in the probability of being enrolled. This is consistent with the enrollment levels we saw before. The difference is likely due to a few factors. Historically, Liberia has had higher levels of education for males than females, as seen by the earlier education trends by gender (Figure 11, Figure 13). It is possible that there is a cultural prioritization of male over female education. 2007 was only four years after the end of the war, and the supply of schools was still far under the demand. It may be the case that with limited opportunities for schooling, people prioritized enrolling their sons over their daughters. There may also be an expectation of girls helping with household duties. Many families were financially struggling after the war, so there may have been increased need for daughters to help with household duties.

The urban variable is also significant and positive. Living in an urban area is associated with a 5.34 percentage-point increase in the probability of being enrolled. This is likely due to the fact that there are more opportunities for enrollment in urban areas, because there are likely more schools and people have to travel a shorter distance to go to school.

Survey: Probit	regression					
Number of stra	ata = 30			Numb	er of obs =	11,466
Number of PSUs	s = 298	Popu	lation size =	1,134,638		
		Desi	gn df =	268		
				F(9,	260) =	121.36
				Prob	> F =	0.0000
enrolled	Coefficient	std. err.	t	P> t	[95% conf.	interval]
female	.0027432	.1036365	0.03	0.979	2013021	.2067885
age	.7680192	.0474127	16.20	0.000	.6746705	.8613679
agesq	0255697	.0018914	-13.52	0.000	0292937	0218457
urban	.0366215	.0816439	0.45	0.654	1241236	.1973665
wealthindex	.3056299	.0382703	7.99	0.000	.2302811	.3809786
deaths	0000156	.0000206	-0.76	0.448	0000561	.0000248
prewar_educ	.060287	.0155841	3.87	0.000	.0296041	.0909699
predgain	.0139277	.0552999	0.25	0.801	0949498	.1228051
educloss	.0013356	.052472	0.03	0.980	1019741	.1046454
_cons	-5.395719	.3502459	-15.41	0.000	-6.085302	-4.706135

Figure 27: Average Marginal Effects for Probit Model

```
Average marginal effects
```

Number	of	strata	=	30	Number of obs =	11,466
Number	of	PSUs	=	298	Population size =	1,134,638
Model V	CE	Linear	riz	ed	Design df =	268

Expression: Pr(enrolled), predict() dy/dx wrt: female age agesq urban wealthindex deaths prewar_educ predgain educloss

	(Delta-method				
	dy/dx	std. err.	t	P> t	[95% conf.	interval]
female	.0007978	.0301307	0.03	0.979	0585253	.0601208
age	.2233498	.0111582	20.02	0.000	.2013808	.2453187
agesq	007436	.0004738	-15.69	0.000	0083689	0065031
urban	.01065	.0236912	0.45	0.653	0359946	.0572946
wealthindex	.0888811	.0108019	8.23	0.000	.0676137	.1101485
deaths	-4.55e-06	6.02e-06	-0.76	0.450	0000164	7.30e-06
prewar_educ	.0175322	.0045469	3.86	0.000	.00858	.0264845
predgain	.0040504	.0161312	0.25	0.802	0277097	.0358104
educloss	.0003884	.0152549	0.03	0.980	0296462	.0304231

Figure 27 is the average marginal effect of the probit model for enrollment rates of people 6-18-years-old in 2007. The explanatory variables in this model that are not significant are female, urban, deaths, predicted education gain (predgain), and education loss (educloss). As I just discussed, the construction of the prewar education, predicted education gain, and education loss variables are driving the insignificance of the female and urban variables.

Deaths due to conflict is not statistically significant. I was treating deaths due to the wars as a measure of the level of destruction in each area, and it is surprising that seems to have no effect. I would have assumed that areas that were most devastated by the wars would have had lower enrollment directly afterwards in 2007. We expect that deaths is highly correlated with the omitted variable aid. It is possible that the negative effects of death and destruction are being essentially cancelled out by the positive effects of aid. I would have expected aid to have more of an impact in 2019, because 2007 was the first year of extremely significant levels of aid. It is possible however that we will see the effects of foreign aid in 2019, and 2007 is picking up the effects of immediate government assistance within Liberia. It is likely the earliest rebuilding efforts were focused on the most devastated areas, so it is possible that there were significant levels of rebuilding for those areas 2003-2007, that we are unable to pick up in the model.

The age variables are statistically significant. Increasing age increases the probability of being enrolled until the turning point of age 15.02, at which point the increasing age begins decreases the probability of being enrolled. This makes sense with the graph of enrollment rates by age.



Figure 8: Percent Enrolled by Age 2007

It seems that the turning point is in fact around age 15, so our model is consistent with this.

The wealth index is also statistically significant and positive. An increase in the factor score of one standard deviation increases the probability of being enrolled by 8.89 percentage-points on average. This implies that people with more household wealth are more likely to be enrolled in school, which makes sense conceptually. People with more wealth are likely better able to afford the costs associated with schooling and have less need for their children to help with household duties.

Prewar education levels are statistically significant. A one-year increase in the average years of education for a given area is associated with a 1.75 percentage point increase in the probability of being enrolled in 2007. Seeing as the mean years of prewar education is 5.12, and the standard deviation is 3.55, this is a significant increase. This implies that the levels of education in an area before the wars began has a significant effect on the likelihood of people in

that same area being enrolled in 2007, holding other factors like wealth and education loss constant. This could be because areas with large amounts of highly educated people before the war had the expectation of their kids being highly education. That expectation was undermined due to the war, but it may have caused the people in those highly educated areas to prioritize getting their kids back into school at the end of the war.

The variable for predicted years of education gain in a counter-factual world without the wars is not significant at any reasonable level of significance. Our best guess based on this sample is that a one-year increase in the years of education that it is predicted an individual would have gained had the war not occurred, is associated with a .45 percentage-point increase in the probability of being enrolled. Because the standard deviation of predicted years of education gain is fairly high (2.05), the 95% confidence interval is also fairly large (-.0277, .0358). This means that we cannot rule out the possibility that predicted education gain has no predictive power, or that is has negative or positive predictive power. We cannot rule out the possibility that an extra year of predicted education gain is associated with up to a 2.77 percentage-point decrease in the probability of being enrolled or that it is associated with up to a 3.58 percentage-point increase in the probability of being enrolled.

The education loss variable is also not significant. Our best guess is that an additional year of education lost due to the war is associated with a .04 percentage point increase in the probability of being enrolled. Just like with the predicted education gain variable, the education loss variable has a fairly large standard deviation (1.99) and a fairly large 95% confidence interval (-.0296, .0304). This means that we cannot rule out the possibility that an additional year of education loss due to the war is associated with up to a 2.96 percentage-point decrease in the likelihood of being enrolled, that it is associated with up to a 3.04 percentage-point increase in

the likelihood of being enrolled, or that it has no effect. Our best guess is that the was very little or no effect, but the 95% confidence interval allows for fairly large deviation from this.

Number of stra	ata = 30			Numb	er of obs	= 11,466		
Number of PSU:	5 = 298			Population size = 1,134,638				
				Desi	gn df =	= 268		
		F(9,	260) =	374.07				
				Prob	> F =	.0000		
				R-so	juared =	e 0.3101		
		Linearized						
enrolled	Coefficient	std. err.	t	P> t	[95% conf.	. interval]		
female	.0049319	.031581	0.16	0.876	0572465	.0671102		
age	.2322968	.0114349	20.31	0.000	.2097831	.2548105		
agesq	0075939	.0004878	-15.57	0.000	0085544	0066334		
urban	.0214514	.0242141	0.89	0.376	0262226	.0691255		
wealthindex	.0864943	.0103995	8.32	0.000	.0660193	.1069694		
deaths	-4.84e-06	6.08e-06	-0.80	0.427	0000168	7.13e-06		
prewar_educ	.0181501	.0045059	4.03	0.000	.0092786	.0270216		
predgain	.0038126	.0166931	0.23	0.820	0290538	.036679		
educloss	.0002465	.0157447	0.02	0.988	0307525	.0312455		
_cons	-1.146812	.0720349	-15.92	0.000	-1.288638	-1.004986		

Figure 28: Linear Probability Model 2007

Survey: Linear regression

Figure 28 is the linear probability model for 2007. As previously discussed, it is important to check that the linear probability model and probit model are similar. The same variables are not significant at the 10% level in the linear probability model that were not significant at the 10% level in the probit model before: female, urban, deaths, predicted education gain, and education loss.

The age variables suggest that an increase in age increases the likelihood of being enrolled until the age of 15.29 and then begins to decrease the likelihood of being enrolled. That is very close to the turning point we found with the probit model (15.02). An increase of standard

deviation of the wealth index factor score is associated with an increase of 8.65 in the probability of being in enrolled. This is extremely close to the increase that the probit model predicted (8.89). A one-year increase in prewar education increases the probability of being enrolled by 1.82 percentage-points, compared to the 1.75 percentage-points that the probit model predicted. The 95% confidence interval for predicted education gain is -.029, .037 compared to -.028, .039 in 2007. The 95% confidence interval for education loss is -.031, .031 compared to -.030, .036 in 2007. Overall, the exact coefficients of the linear probability model were slightly different to the coefficients of the probit model, but they were very close. The coefficients were not different enough to change the overall conclusion of whether or not the variable had a large effect on the probability of enrollment. The same variables were statistically significant, which is encouraging for the validity of the models.

Survey: Probit	t regression					
Number of stra	ata = 30			Numb	er of obs =	14,633
Number of PSUs	s = 325			Рори	lation size =	1,434,310
				Desi	ign df =	295
				F(6,	290) =	52.83
				Prob	> F =	0.0000
encolled	Coefficient	Linearized	+	P> +	[95% conf	intervall
eniorreu	COEFFICIENC	stu. en.	Ľ	FYIC	[95% CONT.	Incerval
female	0365127	.0347063	-1.05	0.294	104816	.0317907
age	.4152536	.040299	10.30	0.000	.3359435	.4945636
agesq	0181838	.0016749	-10.86	0.000	02148	0148876
urban	.0426732	.1093272	0.39	0.697	1724868	.2578332
wealthindex	.3710456	.0443274	8.37	0.000	.2838076	.4582835
deaths	0000469	.0000174	-2.69	0.008	0000812	0000125
_cons	-1.271814	.2330318	-5.46	0.000	-1.73043	8131985

Figure 29: Probit Model Without Education Variables 2019

Figure 30: Average Marginal Effects

Average marginal effects

Number	of	strata	=	30	Number of obs	=	14,633
Number	of	PSUs	=	325	Population size	=	1,434,310
Model \	/CE:	Linear	·17	zed	Design df	=	295

Expression: Pr(enrolled), predict()

dy/dx wrt: female age agesq urban wealthindex deaths

	dy/dx	Delta-method std. err.	t	P> t	[95% conf.	interval]
female	0090656	.0085905	-1.06	0.292	0259719	.0078408
age	.1031014	.0080872	12.75	0.000	.0871854	.1190174
agesq	0045148	.0003367	-13.41	0.000	0051774	0038521
urban	.0105951	.0270008	0.39	0.695	0425435	.0637338
wealthindex	.0921252	.012205	7.55	0.000	.0681053	.1161451
deaths	0000116	4.41e-06	-2.64	0.009	0000203	-2.97e-06

Because of the same issues I discussed earlier, I am running a model without the education variables for 2019 as well. In contrast to 2007, the urban and female variables are not significant at any reasonable level of significance. This implies that both girls and rural areas had significant gains is the rebuilding period 2007-2019, and the gap closed between girls and boys and urban and rural areas. It is likely that the rebuilding programs in Liberia targeted underrepresented groups and areas, and that contributed to girls and rural areas catching up. The urban variable is especially impressive, because it is generally difficult to provide the same amount of schooling in rural areas as urban areas because people are less concentrated, so the schools are generally a further distance from homes.

Survey: Probit regression

Number	of	strata	=	30	Number of obs	=	14,633
Number	of	PSUs	=	325	Population size	=	1,434,310
					Design df	=	295
					F(9, 287)	=	43.05
					Prob > F	=	0.0000

		Linearized				
enrolled	Coefficient	std. err.	t	P> t	[95% conf.	interval]
female	.2903739	.0817959	3.55	0.000	.1293965	.4513513
age	.422776	.0374508	11.29	0.000	.3490714	.4964805
agesq	0185303	.0015416	-12.02	0.000	0215642	0154964
urban	0965444	.1174211	-0.82	0.412	3276336	.1345447
wealthindex	.2969213	.0395182	7.51	0.000	.2191481	.3746946
deaths	0000671	.0000195	-3.45	0.001	0001054	0000288
prewar_educ	.0827622	.0204206	4.05	0.000	.0425737	.1229507
predgain	.0420452	.0313422	1.34	0.181	0196374	.1037278
educloss	0233276	.0294279	-0.79	0.429	0812428	.0345876
_cons	-1.749344	.2124338	-8.23	0.000	-2.167422	-1.331266

Figure 32: Average Marginal Effects

Average marginal effects			
Number of strata = 30	Number of obs	=	14,633
Number of PSUs = 325	Population size	=	1,434,310
Model VCE: Linearized	Design df	=	295
Expression: Pr(enrolled), predict()			

dy/dx wrt: female age agesq urban wealthindex deaths prewar_educ predgain educloss

	Delta-method								
	dy/dx	std. err.	t	P> t	[95% conf.	interval]			
female	.0709165	.020474	3.46	0.001	.0306228	.1112102			
age	.1032524	.0077203	13.37	0.000	.0880584	.1184463			
agesq	0045256	.000317	-14.28	0.000	0051495	0039016			
urban	0235785	.0288922	-0.82	0.415	0804395	.0332824			
wealthindex	.0725155	.0101426	7.15	0.000	.0525545	.0924765			
deaths	0000164	4.91e-06	-3.34	0.001	0000261	-6.72e-06			
prewar_educ	.0202126	.0051297	3.94	0.000	.0101172	.0303079			
predgain	.0102685	.007693	1.33	0.183	0048717	.0254087			
educloss	0056972	.0072052	-0.79	0.430	0198773	.0084829			

Figure 31 is the probit model for the probability of being enrolled in 2019. The only variables that are not statistically significant at the 10% level are urban, education gain, and education loss. The urban variable is not significant due to the education variables. Female is significant, however the education variables are driving this result.

The two age variables are both significant. According to the model, an increase in age is associated with an increase in the probability of being enrolled until age 11.41, then increasing age begins decreasing the probability of being enrolled. This makes sense with the enrollment by age graph for 2019, which is consistent with a turning point of around age 11 and a half.



Figure 9: Percent Enrolled by Age in 2019

This is significantly earlier than the turning point in 2007, which was 15.02. This is likely because directly after the war, people who had missed out on school due to the violence and

school closures enrolled in school. Since many had little or no schooling, they had to enroll in lower education levels, which led to large amounts of people being overage for their school year. In 2019, there are significantly less people starting their education years late because the most education aged people in 2019 did not have the start of their schooling delayed by the war.

The wealth index is statistically significant, and an increase of 1 standard deviation is associated with a 7.25 percentage-point increase in the probability of being enrolled. This is still a large effect, but it is not quite as large as the effect in 2007 (8.89 percentage-points). This implies that wealth, while still important, was slightly less important in 2019 than in 2007. This may be because most people were struggling in 2007, so only fairly wealthy people were able to send their kids to school; however, in 2019, people started to recover financially, so less wealthy people were also able to afford to send their kids to school. It could also be the case that some of the rebuilding or aid programs targeted poorer people and specifically focused on increasing enrollment for them.

Deaths is statistically significant in 2019 at the 5% significance level. However, there is only a .00164 percentage-point decrease for each additional death, which is a 1.64 percentage-point increase for 1,000 additional deaths. That is a fairly significant increase, considering the mean number of deaths is 2,072 and the standard deviation is 2,154.

This means that an increase in deaths does significantly decrease the probability of enrollment, but not at as significantly as you might expect. That implies that the more death an area experienced during the wars, the lower their probability of enrollment is in 2019. This is actually the opposite of what I expected, as I assumed deaths was highly correlated with aid and more deaths would lead to more aid and therefore higher enrollment rates. This variable was not significant in 2007, which means that between 2007 and 2019, kids in areas that experienced lower conflict had their enrollment rates improve more than kids in areas with high-conflict. This is counterintuitive, but it may imply that aid and rebuilding programs were not actually targeting kids in areas with higher conflict levels. It could also mean that areas with higher conflict levels were just not able to recover as much due to the loss in human capital.

An increase in one year of average prewar education in a given area is associated with a 2.02 percentage-point increase in the probability of being enrolled, as compared to an increase of 1.75 percentage-points in 2007. This implies that prewar education slightly more important than it was in 2007. This makes intuitive sense. Historical education rates are an important indicator of current education rates, so it makes sense that it would have a significant effect.

The predicted education gain variable is still not significant in 2019. Our best guess given this sample is that an addition year of predicted education gain is associated with a 1.03 percentage-point increase in the probability of being enrolled. The 95% confidence interval is -.0049, .0254. This means that we cannot rule out the possibility that an addition year of predicted education gain is associated with up to a .49 percentage-point decrease in the probability of being enrolled, that it is associated with up to a 2.54 percentage-point decrease in the probability of being enrolled, or that it has no effect. The 95% confidence interval in 2007 was -.0278, .0358. We cannot determine whether or not there is an effect in 2007 or in 2019, but there is not the possibility of the education gain having as negative of an impact in 2019 as there is in 2007.

Education loss is not significant at the 10% level in 2019, just like it was not in 2007. The best guess given the sample is that a one-year increase in the years of education lost due to the

war is associated with a .57 percentage-point decrease in the probability of being enrolled. The 95% confidence interval is -.0199, .0089. This means that we cannot rule out that a one-year increase in education loss is associated with up to a 1.99 percentage-point decrease in the probability of being enrolled, that it is associated with up to a .89 percentage-point increase in the probability of being enrolled, or that it has no effect. The 95% confidence interval in 2007 was -.0296, .0358.

Based on this sample, we cannot rule out that in 2019, people recovered proportionally to their loss, so that when you control for initial education level, the actual amount of loss is not important. The data are consistent with complete bounce-back, although the 95% confidence interval allows for fairly sizable deviation from this.

Number of stra Number of PSU	ata = 30 s = 325			Num Pop Des F(9 Pro R-s	ber of obs = ulation size = ign df = , 287) = b > F = quared =	14,633 1,434,310 295 47.46 0.0000 0.0935
enrolled	Coefficient	Linearized std. err.	t	P> t	[95% conf.	interval]
female age agesq urban wealthindex deaths prewar_educ predgain educloss _cons	.0766834 .1133148 0049636 0163309 .0685256 0000177 .0209329 .011889 0067609 .0846409	.0237831 .0090441 .0003793 .0317378 .0087026 4.81e-06 .0054072 .0082489 .0075788 .0529097	3.22 12.53 -13.09 -0.51 7.87 -3.67 3.87 1.44 -0.89 1.60	0.001 0.000 0.607 0.000 0.000 0.000 0.151 0.373 0.111	.0298773 .0955156 0057101 0787921 .0513987 0000271 .0102912 0043452 0216762 0194873	.1234895 .131114 0042171 .0461303 .0856526 -8.19e-06 .0315745 .0281231 .0081544 .1887692

Figure 33	: Linear	[.] Probabilitv	<i>Model</i> 2019
1 181110 33	Diffeen	11000001111	1100001 2017

Survey: Linear regression

The same variables are not significant at the 10% significance level in the linear probability model that were not significant in the probit model: urban, education gain, and education loss. An increase in age is associated with an increase in the probability of being enrolled until age 11.41 at which the probability begins decreasing. This is exactly the same turning point that I found in the probit model. An increase of one standard deviation in the factor score of the wealth index is associated with a 6.85 percentage-point increase in the probability of being enrolled compared to 7.25 percentage-points in the probit model. An additional 1,000 deaths is associated with a decrease of 1.77 percentage-points in the probability of being enrolled compared to 1.64 percentage-points in the probit model. An increase in one year of average prewar years of education is associated with an increase of 2.09 percentage-points in the probability of being enrolled, as opposed to 2.02 percentage-point in the probability of being enrolled in the probit model.

The coefficients in the linear probability model are very close to the coefficients in the linear probability model, and the same variables are statistically significant. This implies that the linear probability model is giving a good estimate despite its issues with a dummy dependent variable.

Was there bounce-back?

As discussed in the bounce-back data section, I do not have the data to determine whether or not there was macro-level bounce-back, or complete recovery from the war. I can however use the education loss variable to look at whether or not people recovered proportionately to their level of loss, which I am referring to as micro-level bounce-back. My best guess based off of this sample is that bounce-back did occur; however, I am not able to rule out that bounce-back did not occur, or that a 'more than bounce-back' scenario occurred in which areas with high loss recovered disproportionately well.

In 2007, the 95% confidence interval for education loss implied that a one-year increase in education loss could be associated with up to a 2.96 percentage-point decrease in the probability of being enrolled, that it could be associated with up to a 3.04 percentage-point increase in the probability of being enrolled, or that there was no effect. The best guess based on the sample is that a one-year increase in education loss is associated with a .039 percentage-point increase in the probability of being enrolled, which is approximately no effect. This means that we do not know whether or not bounce-back had occurred by 2007, but our best guess is that it had.

In 2019, the 95% confidence interval for education loss implied that a one-year increase in education loss could be associated with up to a 1.98 percentage-point decrease in the probability of being enrolled, could be associated with up to a .85 percentage-point increase in the probability of being enrolled, or could have no effect. The best guess given the sample is that an extra year of education loss is associated with a .57 percentage-point decrease in the probability of being enrolled, which is not as close to 0 as it was in 2007 but is still a small effect. So, the best guess in 2019 is that people who experienced more loss were not completely recovering proportionately to people who had experienced less loss, but that it was very close. It is possible that this implies that in the longer-term (2019 instead of 2007) people who experienced more loss saw that loss continue into lower enrollment rates; however, it is just not clear with this sample how education loss impacts enrollment levels.

As for macro-level bounce-back, or complete recovery to prewar levels or levels that were predicted in a counterfactual world without the wars, we cannot know for sure that everyone recovered to these levels because of a lack of data; however, the change over time in enrollment rates overall and for various groups imply that at least some bounce-back did occur, even if not complete bounce-back. Overall, enrollment rates increased from 51.01% in 2007 to 81.04% in 2019. This is tremendous growth in enrollment over a 12-year period. Whether or not it was complete or proportional recovery, it is clear that significant recovery did occur in the rebuilding era.

Robustness Checks

Overage Students:

Because so many people were overage for their school year in 2007, I am also going to check what the results for my analyses are for people 16-24 and see if the results are different for those people.

Figure 34: Marginal Effects from the Probit Model in 2007 Without Education Variables

```
Average marginal effects

Number of strata = 30 Number of obs = 5,995

Number of PSUs = 325 Population size = 639,866.74

Model VCE: Linearized Design df = 295

Expression: Pr(enrolled), predict()
```

```
dy/dx wrt: female age agesq urban wealthindex deaths
```

	dy/dx	Delta-method std. err.	t	P> t	[95% conf.	interval]
female	110983	.0171045	-6.49	0.000	1446452	0773207
age	1899539	.057367	-3.31	0.001	3028543	0770534
agesq	.003147	.0014541	2.16	0.031	.0002853	.0060087
urban	.0902163	.0292149	3.09	0.002	.0327203	.1477123
wealthindex	.097413	.0128588	7.58	0.000	.0721063	.1227197
deaths	0000149	6.77e-06	-2.21	0.028	0000283	-1.61e-06

The urban variable is statistically significant. Living in an urban area is associated with a 9.02 percentage-point increase in the likelihood of being enrolled. This is a much greater effect than in the original probit model (5.34 percentage-points). This implies that the people who joined school overage after the war were disproportionately from urban areas.

The female variable in the original probit model for 6-18-year-olds without education variables was statistically significant. Being female was associated with a 7.16 percentage-point decrease in the likelihood of being enrolled. It is statistically significant, and this is a large effect. It implies that males were more likely to join school overage after the war. This has important implications, because the females who did not go back to school overage after the war lost those potential years of education while many of their male counterparts did not. For this specific age cohort, the men are likely to be much more highly educated than females, which could have

effects in the job market, the marriage market, and several aspects of society. This finding is also consistent with the graphs of enrollment by age for each sex.



Figure 35: Enrollment by Age for Females in 2007



Figure 36: Enrollment by Age for Males in 2007

Enrollment rates are generally slightly higher for males than females up until around age 15, but the biggest difference is after age 15. Female enrollment greatly drops off, and male's enrollment, while decreases, maintains relatively high levels.

Figure 37: Marginal Effects of the Overage Probit Model 2007 with Education Variables

```
Average marginal effects
Number of strata = 30
                                                Number of obs =
                                                                      5,095
Number of PSUs
               = 298
                                                Population size = 499,445.75
Model VCE: Linearized
                                                Design df
                                                                =
                                                                        268
Expression: Pr(enrolled), predict()
dy/dx wrt: female age agesq urban wealthindex deaths prewar_educ predgain
           educloss
                         Delta-method
                   dy/dx std. err.
                                         t
                                              P>|t|
                                                        [95% conf. interval]
     female
               -.0914032
                           .0288898
                                      -3.16
                                              0.002
                                                        -.148283
                                                                   -.0345234
               -.1557243
                          .0526361
                                      -2.96
                                              0.003
                                                       -.2593572
                                                                 -.0520914
        age
                                       2.10
      agesq
                .0027735
                          .001321
                                              0.037
                                                        .0001726
                                                                   .0053744
      urban
                 .032136
                         .0293957
                                       1.09
                                              0.275
                                                       -.0257398
                                                                   .0900118
wealthindex
                .0977224
                                       7.71
                                              0.000
                           .0126712
                                                        .0727746
                                                                    .1226701
     deaths
               -.0000145
                          6.68e-06
                                      -2.17
                                              0.031
                                                       -.0000277
                                                                  -1.37e-06
 prewar educ
                .0171842
                           .0042263
                                       4.07
                                              0.000
                                                        .0088632
                                                                    .0255052
   predgain
                                      -3.33
                                              0.001
                -.055117
                          .0165439
                                                       -.0876895
                                                                   -.0225446
                .0583902
   educloss
                            .017082
                                       3.42
                                              0.001
                                                        .0247583
                                                                    .0920221
```

The likelihood of being enrolled decreases with an increase in age for every age in the regression.

Wealth index has a slightly greater impact in this regression as the original one. For every increase in the factor score of one standard deviation there is an increase of 10.42 percentage-points in the probability of being enrolled, as opposed to 8.89 percentage-point increase in the original regression. This implies that household wealth is slightly more important for overage enrollment than for regular enrollment.

The deaths variable is statistically significant at the 5% level unlike in the original model. An addition 1,000 deaths in the area due to the wars is associated with a 1.45 percentage-point decrease in the probability of being enrolled. This implies that the number of deaths in an area is impactful for overage enrollment, but not regular enrollment. It is possible that areas with more destruction and deaths sent their education-aged children back to school at proportional levels to areas with less destruction, but there was a more limited supply of schooling due to the increased destruction in those areas, so overage people had less opportunity to return to school.

The prewar education variable had about the same impact in the overage model as the original one. An increase in one year of prewar education is associated with a 1.70 percentage-point increase in the probability of being enrolled as opposed to 1.75 percentage points in the original. This implies that living in areas of historically high education levels does not matter more for overage people than for regular education-aged people. This is not what I would have expected, as I would have thought that areas with higher levels of education would have been disproportionately more likely to enroll their kids overage.

The predicted years of education gain variable is significant in this overage model, which it was not in the original model. An extra year of predicted education gain is associated with a 5.51 percentage-point decrease in the probability of being enrolled. This is a large impact in the opposite direction of what I expected. The education loss variable is also significant, which it was not in the original model. An increase of one year of education loss is associated with a 5.84 percentage-point increase in the probability of being enrolled. This is also a large impact in the opposite direct than expected. I would have expected that more years of education gain would be associated with a higher probability of being enrolled and more years of education loss would be associated with a lower probability of being enrolled. I would have guessed that these coefficients would have been switched. It is possible that I am getting these coefficients because there is a high correlation between these variables because of how they are constructed.

90

Figure 38: Correlation between Predicted Education Gain and Education Loss for

	predgain	educloss
predgain educloss	1.0000 0.8651	1.0000

16-24 year-olds in 2007

For 16-24 year-olds in 2007, the correlation between the predicted education gain and education loss variables was .8651. Because this is a fairly high correlation, it is possible that the model is having a difficult time differentiating the effects of these two variables, and the coefficients are not an accurate measure of their effects.

Using Different Ranges for Education Loss and Gain Variables:

I was not entirely sure what is the best range to use when creating the education loss and education gain variables. I ended up using ages 36-51 in 2007 and 48-63 in 2019 to estimate the prewar education trend and project the trend to create a counterfactual scenario in which the war never happened. I used 15 years because I wanted to only use the trend directly before the war, but also have enough observations to create a good estimate.

The first civil war began in 1989. There was a coup in 1980 that changed the economic and education trends in Liberia. Because of that, I am going to do a version of the bounce-back variable that only uses those 9 years between 1980-1989. People aged 36 in 2007 and 48 in 2019 were 18 in 1989, so their education was not impacted by the wars. People aged 45 in 2007 and 57 in 2019 were 18 in 1980, so their education was not impacted by anything before the coup.

So, I am going to use the range of 36-45 in 2007 and 48-57 in 2019 for this version of the education loss and predicted education gain variables.

Using this range with the 2007 data, males had an average of -.243 years of education loss as opposed to about a year and females had an average of one year of education loss as opposed about .67 years with the other range. This makes sense because the female education trend was still increasing before the war and the male trend was beginning to decline.

Using the 2019, males had about a year of education loss on average and females had -.164 years of education loss. The values may be different for 2007 and 2019 because there are fewer observations of 48-57 year-olds than 36-45 year-olds.

Figure 39: Marginal Effects for Probit in 2007 with Smaller Years of Education Projection

```
Range
```

Average marginal effects

Number of strata = 30 Number of PSUs = 298 Model VCE: Linearized Number of obs = 11,466 Population size = 1,134,638 Design df = 268

Expression: Pr(enrolled), predict()
dy/dx wrt: female age agesq urban wealthindex deaths prewar_educ predgain
educloss

	Delta-method						
	dy/dx	std. err.	t	P> t	[95% conf.	. interval]	
female	.0031788	.02945	0.11	0.914	054804	.0611615	
age	.2231576	.0111678	19.98	0.000	.2011697	.2451454	
agesq	0074252	.0004742	-15.66	0.000	0083588	0064916	
urban	.0112423	.0224729	0.50	0.617	0330036	.0554882	
wealthindex	.0913968	.0106745	8.56	0.000	.0703802	.1124134	
deaths	-5.81e-06	6.05e-06	-0.96	0.338	0000177	6.11e-06	
prewar_educ	.0178015	.0045131	3.94	0.000	.0089159	.0266871	
predgain	0022775	.0158735	-0.14	0.886	03353	.028975	
educloss	.0079292	.0165704	0.48	0.633	0246956	.040554	

The new range did not change the statistical significance of any of the variables.

With this new range, an addition year of prewar education is associated with a 1.78 percentage-point increase in the probability of being enrolled, compared to a 1.75 percentage-point increase in the model with the original range, which is not a large difference.

Using this range, the predicted education gain variable is not statistically significant, just as it was not significant in the original regression. Based on the 95% confidence interval for this new range, a one-year increase in predicted education gain is associated with up to a 3.35 percentage-point decrease in the probability of being enrolled (2.78 percentage-point decrease in original), associated with a 2.90 percentage point increase in the probability of being enrolled (3.58 percentage-point increase in the original), or no effect.

The education loss variable is not significant at the 10% level using this range, and it was also not significant using the original range. There is a lot of overlap in the 95% confidence intervals for both models. The 95% confidence interval using this new range is -.02470, .04055 and the 95% confidence interval using the original range was -.02771, .03581. Because of this overlap, the model using the new range does not indicate that the range used to create the education variables greatly changes the results. This is encouraging for the validity of the original, because it suggests that regardless of what the optimal range to use is, the results will not be greatly affected.

Figure 40: Marginal Effects for Probit in 2019 with Smaller Years of Education Projection

Range

Average marginal	effects		
Number of strata Number of PSUs	= 30 = 325	Number of obs Population size	= 14,633 = 1,434,310
Model VCE: Linea	rized	Design df	= 295

Expression: Pr(enrolled), predict()

dy/dx wrt: female age agesq urban wealthindex deaths prewar_educ predgain educloss

	Delta-method						
	dy/dx	std. err.	t	P> t	[95% conf.	interval]	
female	.068793	.0211425	3.25	0.001	.0271837	.1104023	
age	.1032731	.0077846	13.27	0.000	.0879526	.1185935	
agesq	0045232	.00032	-14.14	0.000	005153	0038935	
urban	0256303	.0294195	-0.87	0.384	083529	.0322684	
wealthindex	.0740519	.0102167	7.25	0.000	.053945	.0941588	
deaths	0000157	4.89e-06	-3.22	0.001	0000254	-6.12e-06	
prewar_educ	.0205021	.005246	3.91	0.000	.0101777	.0308264	
predgain	.0108326	.0074577	1.45	0.147	0038443	.0255096	
educloss	0113279	.007237	-1.57	0.119	0255705	.0029147	

Using this range in 2019, the same variables are not significant as the ones that were not significant when using the original range: urban, predicted education gain, and education loss.

A one-year increase in prewar years of education is associated with a 2.05 percentagepoint increase in the probability of being enrolled, compared to a 2.02 percentage-point increase in the probability of being enrolled using the original range.

The predicted years of education gain variable is not significant in this model or the model using the original range. The 95% confidence interval for predicted education gain using the new range is -.00384, .02551, while it was -.00487, .02541 using the original range. These are very close confidence intervals.

The education loss variable is not significant using this model or the model with the original range. The 95% confidence interval for the education loss variable is -.02557, .00291. Using the original range, the 95% confidence interval was -.01988, .00848. There is a lot of overlap between the confidence intervals for the two ranges. Just like in the 2007 models, this is encouraging for the validity of my model, because it implies that there is not a huge difference in estimates using different ranges to construct the education variables. So, even if the original 15-year range I decided on is not the optimal range, the results should be fairly similar regardless of the range I use.

Conclusion

The group that benefited the most from the rebuilding process in Liberia in regard to school enrollment rates was females. Education-aged girls have historically had much lower access to schooling than boys in Liberia, and in 2007 they were over 5 percentage-points less likely to be enrolled in school. Through the rebuilding process after the civil wars, girls caught up boys in their enrollment rates, and there were no significant differences in their levels in 2019. Rural areas were also very far behind urban areas in enrollment levels in 2007, with people in rural areas being over 25 percentage-points less likely to be enrolled. By 2019, there was still about a 12 percentage-point difference in the probability of being enrolled for urban vs rural areas, but it is not a significant difference when controlling for other explanatory variables. Areas with lower levels of conflict during the war also recovered more in the post-conflict period than areas with high levels of conflict. This could be due to how the rebuilding programs were implemented, or it could just be due to the loss in human life and capital. The war created

thousands of orphans and displaced thousands of people. It is possible that areas that had higher levels of conflict just struggled to recover from that even with the rebuilding programs.

Other major factors that impacted the probability of kids being enrolled were levels of education in each area before the wars and household wealth. Prewar average levels of education for a given area are important because it shows that historical levels of education are good indicators of current levels of education, even if you control for wealth. This may be due to the expectations for education in the areas. Educated adults may expect their kids to be educated more than less educated adults with the same levels of wealth. Household wealth is important because households with more wealth can afford to send their kids to school and have them not work. It also seems that the rebuilding programs did not aid households with less wealth more than wealthier households.

It is unclear whether or not bounce-back occurred (on a micro or macro level). There was clearly significant recovery in the rebuilding period, because enrollment rates improved greatly both overall and for individual groups. Using the education loss variable as a measure of loss during the war, the model was unable to determine whether or not loss had an effect on enrollment levels, and if so whether these effects were positive or negative. Our best guess based on the sample was that there was little-to-no effect in both 2007 and 2019, which, if true, implies that people recovered proportionately to the loss that they suffered relative to other people in the country; however, the 95% confidence intervals allow for large deviation from this conclusion, as it is also not possible to rule out that education loss had a fairly large positive or negative effect on the probability of being enrolled.

There is space for more research to be done about bounce-back generally and in regard to enrollment. More data would be extremely helpful in attempting to measure bounce-back. To properly measure bounce-back, it would be ideal to have data from before, during, and after the conflict. It is understandably difficult to collect data during and in the direct aftermath of a conflict, but it would be extremely helpful in a bounce-back measure. It would also be helpful to take steps to reduce standard errors. The standard errors for the education variables in my model were fairly high, which made it difficult to determine or even speculate on their effect, because the 95% confidence intervals were quite wide. To reduce standard errors, it would be helpful to increase the sample size. There was already a fairly large sample size in the DHS dataset, so another alternative could be creating a location variable that is smaller than the strata level but larger than the primary sampling unit level. The predicted education gain and education loss variables were both constructed at the strata level because there were not enough observations at the primary sampling unit level to construct them. However, there were only 30 (2019) or 31 (2007) strata, which made the estimates for education by location far less precise. An additional location variable in between the strata and primary sampling unit sizes could have allowed me to construct education variables at more precise levels, but still had enough observations to do so, which could have reduced the standard deviation. A census in the pre-conflict and post-conflict time frames would have been the ideal, because it would have reduced standard deviation by increasing sample size and given more exact location data.

There may be a better method for measuring bounce-back than the method of education loss that I used, and a method to understand bounce-back both at a macro-level, looking at overall recovery, and at a micro-level, looking at relative recovery. Bounce-back in general is an area of post-conflict studies that needs more attention. There is currently not a widely recognized, good way to measure the effects of bounce-back and differentiate them from the effects of other factors such as aid and government institutions. Bounce-back is an important factor to recognize in post-conflict recovery, because if it is ignored other factors like aid could be significantly overestimated.

This model would also be significantly improved with data concerning aid. In my model, aid is an omitted variable. I assume that aid is highly correlated with deaths, but it is very possible that the omitted variable bias from excluding aid is making other estimates in the model biased. There is no dataset for Liberia that measures the amount of aid in the country. An ideal dataset would have an observation for each aid project in Liberia and record the monetary amount of aid given for that project, the location and dates of the project, and the people or sector it was targeting. Including the amount of aid in each location would have significantly improved this model because it would have allowed me to examine the effect of aid, and to control for aid when examining the effects of other variables. For future research on bounce-back specifically, it would be a good idea to choose a case study which has aid data available. It is also important to improve the record keeping and data collection on foreign aid in post-conflict contexts in general. While conflict areas pose significant challenges to data collection, a lot more could be learned about post-conflict societies empirically if more data were available.

References

- Bosede, Awodola. "An Examination of Methods to Reintegrate Former Child Soldiers in Liberia." *Intervention* 10 (March 2012): 30–42. <u>https://doi.org/10.1097</u>.
- Call, Charles, and Vanessa Wyeth, eds. *Building States to Build Peace*. International Peace Institute. Bolder: Lynne Rienner Publishers, 2008.
- Collier, Paul, and Anke Hoeffler. "Aid, Policy and Growth in Post-Conflict Societies." *European Economic Review* 48, no. 5 (October 1, 2004): 1125–45.

https://doi.org/10.1016/j.euroecorev.2003.11.005.

- Cuesta, José, and Ana Abras. "Education and Equal Opportunities among Liberian Children." *Journal of Economic Policy Reform* 16, no. 3 (September 1, 2013): 237–58. https://doi.org/10.1080/17487870.2013.816631.
- Doss, Alan. A Peacekeeper in Africa: Learning from UN Internventions in Other People's Wars. International Peace Institute. Bolder: Lynne Rienner Publishers, 2020.
- Drew, Elizabeth, and Alexander Ramsbotham, eds. "Consolidating Peace: Liberia and Sierra Leone." *Accord: An International Review of Peace Initiatives*, no. 23 (March 2012). <u>https://rc-services-assets.s3.eu-west-1.amazonaws.com/s3fs-</u>

public/Consolidating_peace_Liberia_and_Sierra_Leone_Accord_Issue_23.pdf.

- Econometrics Tutorital. "Logit and Probit: Binary Dependent Variable Models," March 22, 2015. http://econometricstutorial.com/2015/03/logit-probit-binary-dependent-variable-model-stata/.
- "Education | Fact Sheets | Liberia | U.S. Agency for International Development," May 20, 2022. https://www.usaid.gov/liberia/fact-sheets/education.

- Fuest, V. "'This Is the Time to Get in Front': Changing Roles and Opportunities for Women in Liberia." *African Affairs* 107, no. 427 (February 16, 2008): 201–24. https://doi.org/10.1093/afraf/adn003.
- "GDP Growth (Annual %) Liberia | Data." Accessed December 9, 2022. <u>https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?contextual=default&end=2021&lo</u> <u>cations=LR&start=2001&view=chart.</u>
- "GDP per Capita (Current US\$) Liberia | Data." Accessed November 20, 2022. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=LR.
- Guimbert, Stephane, Keiko Miwa, and Duc Thanh Nguyen. "Back to School in Afghanistan:
 Determinants of School Enrollment." *International Journal of Educational Development* 28, no. 4 (July 1, 2008): 419–34. <u>https://doi.org/10.1016/j.ijedudev.2007.11.004</u>.
- Hegre, Håvard, Gudrun Østby, and Clionadh Raleigh. "Poverty and Civil War Events." *Journal of Conflict Resolution* 53, no. 4 (August 2009): 598–623.

https://doi.org/10.1177/0022002709336459.

- Kruk, Margaret E, Peter C Rockers, Elizabeth H Williams, S Tornorlah Varpilah, Rose Macauley, Geetor Saydee, and Sandro Galea. "Availability of Essential Health Services in Post-Conflict Liberia." *Bulletin of the World Health Organization* 88, no. 7 (July 1, 2010): 527–34.
 https://doi.org/10.2471/BLT.09.071068.
- Kumar, Ajitesh. "Logit vs Probit Models: Differences, Examples." *Data Analytics* (blog), January 15, 2023. <u>https://vitalflux.com/logit-vs-probit-models-differences-examples/</u>.
- Lai, Brian, and Clayton Thyne. "The Effect of Civil War on Education, 1980—97." *Journal of Peace Research* 44, no. 3 (May 2007): 259–376. <u>https://doi.org/10.1177/0022343307076631</u>.

Let Girls Lead. "Policy Brief: Liberia Girls' Education Policy." Accessed December 7, 2022. https://search.issuelab.org/resource/liberia-girls-education-policy.html.

- Levey, Elizabeth J., Claire E. Oppenheim, Brittany C. L. Lange, Naomi S. Plasky, Benjamin L.
 Harris, G. Gondah Lekpeh, Isaac Kekulah, David C. Henderson, and Christina P. C. Borba. "A
 Qualitative Analysis of Factors Impacting Resilience among Youth in Post-Conflict Liberia." *Child and Adolescent Psychiatry and Mental Health* 10, no. 1 (August 12, 2016): 26.
 https://doi.org/10.1186/s13034-016-0114-7.
- "Liberia Joint IDA-IMF Staff Advisory Note on the Poverty Reduction Strategy Paper : First Annual Progress Report." Text/HTML. World Bank, June 9, 2010.

https://documents.worldbank.org/en/publication/documents-

- reports/documentdetail/802981468272720778/Liberia-Joint-IDA-IMF-staff-advisory-note-onthe-poverty-reduction-strategy-paper-first-annual-progress-report.
- "Liberia National Education Profile 2018 Update." Education Policy and Data Center, 2018. https://www.epdc.org/sites/default/files/documents/EPDC_NEP_2018_Liberia.pdf.
- Liu, Shelley X. "How War-Related Deprivation Affects Political Participation: Evidence from Education Loss in Liberia." *Journal of Peace Research* 59, no. 3 (2022): 353–66. https://doi.org/10.1177/00223433211019460.
- Momodu, Samuel. "Second Liberian Civil War (1999-2003) •," October 3, 2017. https://www.blackpast.org/global-african-history/second-liberian-civil-war-1999-2003/.
- Nworgu, Boniface G, and Loretta N Nworgu. "Urban-Rural Disparities in Achievement at the Basic Education Level: The Plight of the Rural Child in a Developing Country." *Developing Country Studies* 3, no. 14 (2013): 128–40.

- Parks, Bradley C, and Nakul Kadaba. "The Elusive Quest for Effective Aid Management in Liberia." Aid Data, October 20, 2010. <u>https://www.aiddata.org/blog/the-elusive-quest-for-effective-aid-management-in-liberia</u>.
- Pedace, Roberto. "3 Main Linear Probability Model (LPM) Problems." For Dummies, March 26, 2016. <u>https://www.dummies.com/article/business-careers-money/business/economics/3-main-linear-probability-model-lpm-problems-156466/</u>.
- Relief Web. "Gates Foundation, Partners Pledge \$90 Million to Boost Incomes of Small Farmers in Africa," February 19, 2009. <u>https://reliefweb.int/report/cameroon/gates-foundation-partners-pledge-90-million-boost-incomes-small-farmers-africa</u>.
- Republic of Liberia: Ministry of Internal Affairs. "Overview of Liberia," n.d. <u>https://www.mia.gov.lr/2content.php?sub=210&related=40&third=210&pg=sp</u>.
- *Reuters*. "Timeline: Liberia: From Civil War Chaos to Fragile Hope." November 7, 2011, sec. World News. <u>https://www.reuters.com/article/us-liberia-election-events-idUSTRE7A62BN20111107</u>.
- Roessler, Philip, and Rob Blair. "Chinese Development Finance and the African State: A Capacity-Legitimacy Trade-Off?" Aid Data, June 19, 2014. <u>https://www.aiddata.org/blog/chinese-</u> <u>development-finance-and-the-african-state-a-capacity-legitimacy-trade-off</u>.
- St. Louis Federal Reserve. "Real GDP at Constant National Prices for Liberia." Federal Reserve Economic Data, 2023. <u>https://fred.stlouisfed.org/series/RGDPNALRA666NRUG</u>.
- "The DHS Program Country Main." Accessed November 20, 2022.

https://dhsprogram.com/Countries/Country-

Main.cfm?ctry_id=22&c=Liberia&Country=Liberia&cn=&r=1.

The Economist. "Aid Brought Liberia Back from the Brink." June 29, 2017. https://www.economist.com/international/2017/06/29/aid-brought-liberia-back-from-the-brink. The International Bank for Reconstruction and Development, and The World Bank. "Liberia Education Country Status Report: Out of the Ashes - Learning Lessons from the Past to Guide Education Recovery in Liberia," December 2010.

https://documents1.worldbank.org/curated/en/257321468057236139/pdf/636350WP0P117600Li beria0CSR000Edit4.pdf.

The World Bank. "Population, Total - Liberia," 2023.

https://data.worldbank.org/indicator/SP.POP.TOTL?locations=LR.

- "Total Official Development Financing ODF | OECD International Development Statistics | OECD ILibrary." Accessed December 9, 2022. <u>https://www.oecd-ilibrary.org/development/data/oecd-</u> international-development-statistics/total-official-development-financing-odf_data-00071-en.
- Waydon, Emmanuel Boone, Liu Ying, and Barbara L Ketter. "FREE AND COMPULSORY

PRIMARY EDUCATION POLICY IN LIBERIA: GAP BETWEEN PROMISE AND ACTUAL PERFORMANCE." *Educational Research International* 5, no. 1 (February 2016): 17.

Werker, Eric, and Jasmina Beganovic. "Liberia: A Case Study," 2011.

https://www.theigc.org/sites/default/files/2011/07/Werker-2011-Liberia-Case-Study.pdf.

- Wooldridge, Jeffrey. Introductory Economics: A Modern Approach. 5th ed. South-Westeren, Cengage Learning, 2012. <u>https://jaimedv.com/eco/4c1-ecomet/jeffrey-m-wooldridge--econometrics--</u> <u>book.pdf</u>.
- World Bank. "Donor Coordination." Text/HTML, September 30, 2022. https://www.worldbank.org/en/country/liberia/overview.
- "World Bank Maps." Accessed November 20, 2022.

https://maps.worldbank.org/projects/wb/sector/Education/country/LR/Liberia?active=1.