March 2023

Natural Disasters and Human Capital: Empirical Evidence from Indonesia

Lei Lv
University of South Florida

Natural Disasters and Human Capital: Empirical Evidence from Indonesia

by

Lei Lv

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Economics
College of Arts and Science
University of South Florida

Major Professor: Giulia La Mattina, Ph.D.
Padmaja Ayyagari, Ph.D.
Benjamin Jacob, Ph.D.
Xin Jin, Ph.D.
Gabriel Picone, Ph.D.

Date of Approval:
March 22, 2023

Keywords: Earthquake,Water Safety, Child Marriage, Education,

ACKNOWLEDGMENTS

I would never be able to do this without the support from my committee. My major professor, Dr. Giulia la Mattina, met me when I had no research skills or background. She helped me build my skills and knowledge today and showed me how to be a good researcher: always be curious about the world and persistent when faced with challenges. Dr. Padmaja Ayyagari is always available and attentive to every question and request I had. Dr. Xin Jin, Dr. Gabriel Picone, and Dr. Benjamin Jacob provided extensive support and knowledge to support my research. Many other professors, seminar speakers, and colleagues gave me valuable comments and feedback. I would never have done this without their support.

Aside from my professors, I want to thank my family and friends. There is a special place for Tyler Faust, who shed light on me during my darkest time and gave me hope for a better life. Thanks to Sai Yang and Michael Wright, who accompanied me during the most difficult time and showed me a wonderful life to live. Thanks to Minglu Sun and Ricky Zhang for being my friend and keeping me happy. Thanks to my mom and daddy, who supported me financially and emotionally throughout the process. And many thanks to whoever sparked wonders into my life and made me who I am today.

Lastly, I also want to thank myself for not giving up and did not end my life during the most depressing time. To myself in 2016: you made it, you are happy, and life is amazing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In developing countries, natural disasters could destroy physical capital and adversely affect human capital accumulation by disrupting individual decisions. Such decisions play a critical role in determining individuals' human capital accumulation process and have a lifelong effect on their happiness and economic prosperity. To better understand how natural disasters affect human capital in developing countries, this dissertation uses the earthquakes in Indonesia as a natural experiment to study how this earthquake affects health, child marriage, and education. For the first chapter, I study how the 2006 Yogyakarta Earthquake a affects water-related acute disease symptomsin the short and long run. By tracking individuals before and after the earthquake, I identify the wateidentifiedrborne diseases related symptoms decreased significantly one year and eight years after the earthquake. The improved access to safe water, which is the major concentration of the reconstruction program could explain t,he change, showing the that robust reconstruction program could turn a disastrous event into a beneficial one. For the second chapter, I explore how natural disasters affect the hazard rate of being married befobefore the age ofin Indonesia from 1990 to 2014. By tracking the migration and earthquake overtime, my coauthors and I show that earthquake resulted in less girls married before age of 18 in the rural area, more girl married in the urban area. This result indicates a graving and pressing needs for local and international governments to implement policies to alleviate negative impof from the natural disasters. For the last chapter, I explore how the 2006 Yogyakarta affected the total education years and likelihood of finishing different school level in the earthquake affected area.

The result indicates that education process was largely disrupted after the earthquake, leading to education years lost and less students were able to finish senior high school and this effect is more pronounced among boys. Evidence from the labor market indicates that more boys left school for work and higher paid reconstruction jobs, giving evidence that reconstruction programs might have attracted boys to leave school and forced them to less education as a result. Such result indicates that although reconstruction might have helped the affected area, but it also caused unwanted consequences in education and local government might needs to provide better education policies to the affected region for long term human capital development.

CHAPTER ONE: NATURAL DISASTERS, HEALTH, AND WATER SAFETY

With increasing seismic and meteorological hazard incidences in developing countries over the past decades, natural disasters have imposed drastic risks and adverse effects on the health conditions of the affected population (United Nations, 2020; Arya & Agarwal, 2012). Health is a critical factor for human capital, deciding the productivity and welfare of individuals and community welfare and regional economic development (Bloom et al., 2004; Ogundari and Awokuse, 2018; Thomas and Frankenberg, 2002). The negative impacts of natural disasters also cause individuals to lose their productivity, face asset damage, and suffer from welfare loss, and those changes lead to long-term adverse effects on their happiness, productivity, and even their children (Caruso and Miller, 2015; Caruso, 2017). Among all risk factors, access to clean and safe water is critical to health conditions after natural disasters (Popkin, D'Anci & Rosenberg, 2010). As the affected populations are likely to drink and use unsafe water resources or to be displaced to other locations, the absence of safe water could result in different acute infections such as skin and soft tissue infections and gastrointestinal infections and damage individual health conditions from loss of health and cost of medical services (Ho et al., 2019; Kouadio et al., 2012; Watson et al., 2007). The existing literature confirms that developing countries face a higher risk of water safety issues after natural disasters. The risk factors include the underdeveloped water supply system (Tsoukalas and Tsitsifli, 2018), unable to repair damaged water processing facilities on time (Omar et al., 2017), limited safe water from other locations due to damaged transportation (Watson et al., 2007), and heavily rely on the natural water resources (Kasozi et al., 2019). Nevertheless, robust assistant programs after natural disasters could help to lower the risk of infectious diseases and promote better health conditions (Rundblad et al., 2010; Watson et al., 2007).

However, the existing research mainly has focused on the immediate effect of natural disasters on water safety and overlooked the long-term and persistent impacts of natural disasters on health and water safety. Do those negative impacts from natural disasters on water safety continue years after natural disasters, or do they diminish after a specific time? On top of that, as many affected regions receive extensive assistance after natural disasters, could the positive effect coming from the reconstruction programs offset the negative effect of the natural disasters both in the short term and long term? I used a large-scale, longitudinal household survey in Indonesia and combined geographic information system (GIS) data to answer these questions. By mapping the effect of the 2006 Yogyakarta Earthquake on water safety one year and eight years after the earthquake, I explore how a destructive earthquake affects health conditions and waterborne diseases with an extensive reconstruction program in the affected area.

A 6.3 magnitude earthquake happened in the densely populated provinces of Central Java and Yogyakarta on May 27, 2006. The earthquake resulted in more than 5,700 deaths, injured about 40,000 people, and destroyed around 350,000 homes, causing a $3.1 billion loss and marking the most costly natural disaster in developing countries (Java Reconstruction Fund, 2012; Kirchberger, 2017). Because of the poorly constructed buildings in the affected area and the earthquake shaking intensity amplified by a relatively shallow earthquake depth, the affected area faced much more devasted damage and loss than previously expected. Additional to the earthquake damage, the already weak water supply system worsened the water safety condition. In 2014, about 18% of households relied on surface water for drinking, and only 11% of households had access to pipeline water in Indonesia, conditioning in poor water quality and regular supply interruption (Statistics Indonesia, 2014). Given Indonesia's weak water and sanitation conditions, diarrhea remained one of the fatal health concerns, resulting in 31% of the post-neonatal death and 25% of child mortality (UNICEF, 2012). After the 2006 Yogyakarta earthquake, the earthquake damaged the water supply stations leading to massive temporary water supply disruption within the affected area (Java Reconstruction Fund, 2012). Disrupted water resources could result in diarrheal infections such as Typhoid and Gastritis, Hepatitis infections such as Hepatitis A and Hepatitis E, and external infections such as Leptospirosis and Cellulitis (WHO, 2006). The affected area also received considerable international and domestic assistance for post-disaster reconstruction to mitigate the negative impacts (Java Reconstruction Fund, 2012).

The Java Reconstruction Fund (JRF) was the reconstruction program led by the local government and guided by the United Nations to restore the earthquake-affected area (Java Reconstruction Fund, 2012). JRF was fully allocated a total of US$ 89.91 million from different resources to finance five projects in the affected area to conduct post-disaster reconstruction of the communities and improve the general livelihoods of the population. Of the five projects, Community Settlement Reconstruction and Rehabilitation Project (CSRRP) took US$ 71.26 million, counting about 80% of the total expenditure. CSRRP aimed to satisfy the housing needs arising after the earthquake and restore other community facilities, including village roads and footpaths, retaining walls, water supply, and sanitation facilities. As of June 30th, 2011, the program restored 400 water supply and sanitation facilities. As a result, many affected populations had access to water in their houses rather than using natural water resources or sharing water supply in communities. Thus, although WHO (2007) reported disease outbreaks immediately after the earthquake, the JRF could have helped improve the water safety condition and reduced the infection risk in the affected area. But whether the negative impacts from the earthquake are more significant than the positive impacts from JRF on water safety or the adverse effect is less than the improvement from JRF, the answer is unclear.

This paper studies the short-term and long-term effects of the earthquake on health through water safety using geographic information system (GIS) tools to map out the earthquake's impact. I found evidence that waterborne disease symptoms significantly decreased one year and eight years after the earthquake, but non-waterborne disease symptoms did not fall.[1] Thus, decreased waterborne disease risk might have improved the reported health condition of the affected population increased one year after the earthquake, but not in the long term. Although this study cannot verify the actual reason why the reported health did not improve in the long term, one explanation is that since an acute infection does not reduce overall health on a large scale, people do not feel that they are healthier in the long term once they have used to a condition where fewer infections. Moreover, my results also show that households had better access to safe drinking and using water, which are the focuses of the reconstruction program, and the general living environment as a whole also improved both in the short term and long term. Those changes in the household level could explain why the waterborne disease symptom incidence decreased significantly in the short term and long term, proving that the reconstruction programs were successful and provided long-term positive effects in the affected region. A simple Cost-Effectiveness Analysis (CEA) indicates that the reconstruction program provided an effective solution in managing waterborne diseases compared to other large population programs, but still less effective than programs at small scale targeting waterborne diseases.

[1] The main dataset. Indonesia Family and Life Survey (IFLS) only gives symptoms not diseases incidence. Thus, I do not know what type of diseases the individual acquired, only symptoms.

Thus, this study suggests that reconstruction after the earthquake could help reduce the incidence of waterborne diseases and improve the living environment, both in the short and long term. It also provides strong empirical evidence to focus on rebuilding and restoring water and sanitary facilities for after-disaster management in developing countries.

This paper contributes to the current economics literature by expanding on the effect of natural disasters on health outcomes. I add to the literature by exploring how natural disasters affect individual welfare. Natural disasters could spur poverty level due to loss of property and livelihood (Khayyam, 2020), affect marriage decisions and increase child marriages (Ayyagari et al., 2022; Corno et al., 2020), generate education and health loss (Caruso and Miller, 2015; Caruso, 2017; Lv, 2022), promote higher wage growth (Kirchberger, 2017), and increase fertility and reduce child spacing (Nandi et al., 2017). To build on the literature, I explore the effect of natural disasters on health and water safety. Natural disasters undermine the health condition by introducing more diseases outbreak afterward (Kouadio et al., 2012; Watson et al., 2007). My paper suggests that effective and robust water safety programs could help mitigate those negative impacts, control disease outbreaks, and improve personal health in the long run. In addition, to my knowledge, this is the first study to explore the long-term causal effect of natural disasters on water safety and health in developing countries when large reconstruction programs were also implemented after major natural disasters.

The remainder of this paper is as follows. Section 2 presents the datasets used in this paper and defines variables. The research design and statistical analysis of the individual and household are described in section 3. Section 4 presents the results of observations living in the same communities over time, and section 5 explores those who moved after the earthquake. Section 6 checks the threat to the identification and provides robustness checks for the main result. Section 7 conducts a Cost-Effectiveness Analysis (CEA) to understand the reconstruction program's overall effectiveness better. Finally, a discussion of the results and policy implication is in section 8. section 9 concludes this paper.

Data

I use three datasets for this study: the Indonesian Family Life Survey (IFLS) for individual, household, and community panel information and the United States Geological Survey (USGS) ShakeMap for the earthquake ground shaking intensity. One key research aspect is how the earthquake reconstruction could have altered the effect of the earthquake on water safety and resulted in potential benefits for the affected areas. The spatial distribution data on the reconstruction resources could have helped me identify this potential effect accurately. Unfortunately, no data provides information on the geographic distribution of the reconstruction resources.

Indonesian Family Life Survey (IFLS)

I use the Indonesian Family Life Survey (IFLS) data to extract personal information, including the incidence of acute disease symptoms, age, residential location, and household living environment. The Indonesian Family Life Survey (IFLS) is an ongoing longitudinal survey in Indonesia that started in 1993. It records the respondents of their individual, household, and community characteristics by the survey interviewers. The IFLS represents about 83% of the Indonesian population and contains over 30,000 individuals living in 13 of the 27 provinces from 1993 to 2014/2015. The first wave was conducted in 1993, 2^{nd} wave in 1997, 3^{rd} wave in 2000, 4^{th} in 2007/2008, and the last wave in 2014/2015. With the richness of the IFLS, I create the incidence of acute disease symptoms at the individual level and their household living environment by the time of the survey.[2]

To identify the effect of the earthquake and its reconstruction on waterborne diseases, I extend the scope of this research and include non-waterborne diseases. Thus, I create two groups of acute incidences: waterborne and none.[3] Waterborne incidence includes stomachache, diarrhea, vomiting, skin infection, and eye infection, and none-waterborne incidence includes coughing, headache, running nose, difficulty breathing, and fever. I define waterborne incidence as it could be directly related to waterborne infections such as virus or bacteria-induced gastrointestinal, cutaneous, and out-layer body tissue infections. I exclude headache and fever in the main analysis as other infections, such as upper-track respiratory infections, could also lead to those two symptoms. But I add them back to the robustness check section to count for potential bias. Furthermore, to understand why symptoms have changed after the earthquake, I create a vector of household environment variables for each household, including access to safe drinking water, indoor water access, access to safe using water, access to toilets, access to sewage, and access to garage service. In addition, IFLS also provides community-level information with exact GPS location; this data helps to identify the effect of the earthquake at the community level. All variables and definitions used for this paper are demonstrated in Appendix A.

I use IFLS3(2000), IFLS4 (2007/2008), and IFLS5(2014/2015) as the study period and include IFLS2 (1997) to test the identification assumption in the later section. As most of the

---

[2] IFLS also provides information for chronic diseases, but it is out of the scope of this study.
[3] In public health, incidence means any symptoms appeared within a certain time frame and they are considered as new occurrence within that period. Prevalence is the proportion of persons in a population who have a particular disease or attribute at a specified point in time or over a specified period. For example, new poverty population is similar as incidence and total poverty population is prevalence.

reconstruction finished in 2009, IFLS4 records the short-term effect of the earthquake or during the reconstruction period by comparing the individual and households between IFLS3 and IFLS4. Meanwhile, IFLS5 describes the conditions of the individual after the reconstruction is finished, showing the long-term effect of the earthquake. By tracking the same individuals and comparing the changes between IFLS3 and IFLS5, I identified the long-term impact of the earthquake on the affected individuals and households. Thus, together with three different waves, I explored the effect of the earthquake with a robust reconstruction program on the incidences of acute symptoms and the accessibility to safe environments among affected households both in the short and long term.

As I study the effect of changes in the environment to explain the differences in waterborne disease incidence after the earthquake, individuals' residential location is critical in eliminating the confounding effects in estimation. To limit the confounding factors, I divided my study samples into two groups: never-mover and ever-mover. Never-movers are those households who lived within the same community over the study period. If a household moved within the same communities, such as between different villages, I still counted them as never-movers as I do not have information for such movement.[4] Never-movers are those who moved to other communities after IFLS3 at the subdistrict or higher administrative level.[5] If someone moved after IFLS4, they are defined as a never-mover when studying the short-term effect (from IFLS3 to IFLS4) and an ever-mover when looking at the long-term impact (IFLS3 to IFLS5). I

[4] As moving between villages are not publicly available, migration between communities is the most accurate information for migration from IFLS.

[5] IFLS has village, subdistrict, community, district, and province administrative level. Village, subdistrict, district, and province are recoded by BPS code, which is a code used by the local government similar to zip codes. Community is a IFLS survey level, and it is between subdistrict, and district. Moving between subdistrict could either moving within a community or out of community. Such any movement at the subdistrict level counts as moving outside of community.

do not have information on the location of the movement at the household level, giving the survey collects data on whether the household has moved since the last survey. Although detailed individual-level migration history is available, it does not inform if individuals moved out of the community. Given that the earthquake information is at the community level, using personal migration history would deviate from the current analysis and require more detailed earthquake information and migration information at different administrative levels. Thus, I only consider migration at the household level, not personal, and assume that individuals are moving with their households.

United States Geological Survey (USGS)

Indonesia is a country with constant seismic risk. Figure 1 shows earthquakes that happened onshore with a magnitude higher than 5.5 from 1980 to 2015 in Indonesia. For this study, I choose one of the most costly earthquakes in the developing world, the 2006 Yogyakarta earthquake. To study the effect of this earthquake, I used ShakeMap Modified Mercalli Intensity (MMI) from the US Geological Survey. Each earthquake has an associated earthquake magnitude on the Richter scale, indicating the total energy released by the earthquake. However, the modified Mercalli intensity (MMI) provides a more accurate measurement of the actual effect of the earthquake as it measures the ground shaking intensity and how likely it would result in damage to the affected area (USGS, 2021). MMI composes increasing intensity ranging from imperceptible shaking to catastrophic destruction by a Roman numeral from I to XII. Compared to traditional measurements such as distance to epicenter or house damaged per capita, MMI isolates the geographic heterogeneity such as spaces, soil types, and earthquake types to accurately identify the true impact of the earthquake at specific locations (Zhao et al., 2006). Figure 2 shows the Modified Mercalli Intensity in Java Island recorded by the earthquake sensor machines. ShakeMap provides MMI data increase with 0.2 intensity units, in which darker red indicates a higher level of MMI, lighter red shows a lower intensity, and the white color presents no recorded intensity or effect in that location.[6] I spatially linked the IFLS communities with MMI shapefile data using their GPS locations to assign the MMI at the community level.[7]

Furthermore, considering the most affected area, Yogyakarta City, is a more developed area than other parts of Indonesia, I follow Kirchberger (2017) and select communities centered within 50 kilometers of cities with more than 100,000 population in 2000 to construct a more homogenous sample. Including samples from the rural or less underdeveloped areas as a control group might bias the effect of the earthquake and lead to confounding effects in estimation. Therefore, I select the control group location near major cities. Table 1 shows the health posts difference and other communal differences between selected and excluded communities. The community-level data indicates that the excluded communities have far worse living conditions than the communities selected. Figure 2 illustrates the communities chosen for this study. The red dots indicate that those communities are located within the at least 50 kilometers buffer of large cities within Java Island. The white dots are communities excluded from my study.

To identify the earthquake MMI at the community level, I use ArcGIS Pro spatially links the IFLS communities with the earthquake shapefile from USGS. This process gives each of the community GPS locations one earthquake MMI level. I further linked the earthquake MMI level to individuals and households based on their community residential locations at the time of the earthquake. Table 2 summarizes the data used for this study.

[6] Due to the technology limitation, intensity lower than 2.8 are generally not recorded by the sensor machines.
[7] The individual resident location at the village level is not accessible due to the privacy concerns.

Empirical Strategy

Individual Fixed Effect

Using the 2006 Yogyakarta Earthquake as the natural experiment and drawing data from IFLS, I use the individual fixed effect model to identify the causal effect of the earthquake on acuate disease incidences and reported health conditions. Assuming that in the absence of the earthquake, trends in the disease symptoms across individuals would have been the same, this study design employs a strategy of within-individual comparison along two dimensions: one across geographic location and one across time (Brown and Velásquez; 2017). When compared across time, I compare the incidence of symptoms of the same individual before and after the earthquake, controlling time-invariant characteristics which cannot be affected by the earthquake. When compared across space, I explore how different earthquake intensities affect the incidence of symptoms across different communities affected by the earthquake or not. The treatment equals to 1 if the individual or household experienced an MMI higher than higher or equal to 6 and 0 otherwise. I define an earthquake as MMI higher or equal to 6 since the earthquake would damage resistant and vulnerable buildings (Wald et al. 1999). A robustness check using MMI 5 and MMI 5.5 to define earthquakes is provided in the later section. This research design helps control omitted variables at the individual level and isolates the direct causal impact of the earthquake on the symptom incidences. Specifically, I estimate the following model:

$$Y_{ics} = \beta_0 + \beta_1 M_{ics} + \beta_2 Age + \beta_3 AgeSquare + \mu_t + \theta_i + \vartheta_s + \varepsilon_{isc}$$

Where Y_{ics} is the incidence of symptoms, which equals 1 if individual i experienced any specific types of symptoms over the past four weeks and lived in the community c at survey wave s. M_{ics} is the earthquake indicator and equals to 1 if the MMI is higher or equal to 6 at that community, and 0 otherwise. *Age* and *AgeSquare* record age's linear and nonlinear effect on the incidence of symptoms over time.[8] μ_t is the month-fixed effect recording the unobservable effect from the seasonality of diseases (Azage et al., 2017). θ_i is the individual fixed effect tracking the unobservable unique characters, which could have affected the symptoms' incidences of the symptoms, and ϑs is the survey wave fixed effect demonstrating different levels of symptoms reported across surveys. My coefficient of interest is the β_1 recording the causal effect of the earthquake on the incidence of a vector of disease symptoms. The underlying assumption is that the individuals before the earthquake did not have different trends of disease incidence in affected and non-affected communities. This assumption is verified in section 6. The standard errors are clustered at the community level to allow serial correlation within the same geographic units.

Household Fixed Effect

Next, I explore the mechanism of the disease symptom incidence changes based on the household living environment factors. By examining the causal effect of the earthquake on the household living environment, I use the household living environment change to explain the differences in the incidence of disease symptoms among individuals and estimate the effect of the reconstruction program. Those environments include access to safe drinking/using water, in-house water supply, sewage system, toilets, and garbage processing service. Similar to the individual fixed effect, this estimation compares the accessibility to a safe environment before and after the earthquake for households living in earthquake-affected and not-affected communities. The following specification demonstrates the estimation:

[8] To ensure the strictly exogeneity of the estimation, no other control variables are used for this analysis since other variables such as smoking behavior could be affected by the earthquake which explains the incidences of the diseases at the same time.

$$Y_{hcs} = \beta_0 + \beta_1 M_{hcs} + \theta_h + \vartheta_s + \varepsilon_{isc}$$

Y_{hcs} indicates if the household h can access a specific household living environment factor living in community c during survey wave s. It takes the linear value 1 if the household access to a particular type of safe living environment factor, and 0 otherwise. M_{hcs} is the earthquake indicator equal to 1 if the MMI is higher or equal to 6, zero otherwise. θ_h is the household fixed effect tracking the unobservable household characters that could have affected accessibility to safe living environment factors, and ϑ_s is the survey wave fixed effect recording the difference between different survey waves. β_1 shows the causal impact of the earthquake on the accessibility to a safe environment if the household did not experience different levels of the living environment factor changes in the affected and non-affected communities before the earthquake. Section 6 verifies the assumption using a placebo test between IFLS2 and IFLS3. The standard errors of this estimation are clustered at the community level to allow serial correlation within the same communities.

Multiple Hypothesis Testing

Additionally, since I studied a vector of disease symptoms and living environment factors, I created an index recording all disease incidences and living environments to improve the statistical power of the estimation by following Currie et al. (2020) and Kling et al. (2007). As my outcome variables orient in the same direction showing the same effect: a higher value among diseases incidence means a higher likelihood of experiencing diseases symptoms, and a higher value among living environment means a higher probability of accessing a better environment factor, I aggregate the same direction outcomes within a domain as indices, Y. This method improves statistical power to detect effects of the earthquake over a vector of outcomes. Mathematically, the indices Y is defined as the equally-weighted average of z-scores of its components. A larger index score indicates a more adverse effect on disease incidence and a more beneficial outcome for living environments. The z-scores are calculated as follows:

$$z_{ics} = \sum \frac{X_{ics} - \overline{X_{ics}}}{sd_{ics}} \quad or \quad z_{hcs} = \sum \frac{X_{hcs} - \overline{X_{hcs}}}{sd_{hcs}}$$

I estimated indices Y with a summation of equally-weighted z-scores using treatment outcome, subtracted the control group average first, and divided by the control group standard deviation with each component of individuals i or households h. The control group is defined as those individuals or households who were not affected by the earthquake. Therefore, based on the specifications, the control group variables take a mean of 0 and a specific standard deviation for each disease symptom and living environment component.

Standard statistical techniques could over-rejection null hypotheses when considering multiple hypothesis tests simultaneously. Therefore, I also calculate the step-down adjusted p-values to correct the multiple hypothesis testing by following Romano and Wolf (2005). By setting a list of binary decisions concerning all individual null hypotheses, this procedure constructed a better statistical estimation using a stepwise multiple-testing procedure that asymptotically controls the familywise error rate. As this procedure considers the probability of rejecting at least one true null hypothesis in a family of hypotheses under the test, the overall results are more conservative than the original results. Both the indices and Romano-Wolf p-values are reported in the result section.

Results

Disease Incidence of Never-movers

First, I explore the effect of the earthquake on waterborne diseases and their related symptoms. The estimation results are presented in Table 3. By comparing IFLS3 and IFLS4, the result shows that skin infection and eye infection decreased significantly among individuals who lived in non-earthquake communities, but no significant change in stomachache, vomiting, and diarrhea. Skin infection increases by 8.45 percentage points, and eye infection decreases by 4.98 percentage points, counting for 81% and 93%decrease from the group average. Furthermore, the waterborne diseases index indicates a substantial reduction among earthquake-affected individuals compared to unaffected individuals. Thus, waterborne disease incidence decreased significantly in the affected areas one year after the earthquake, and Romano-Wolf P-values indicate that the effect of skin and eye infection is not likely to be random at 0.01 level. This result differs from most previous studies on natural disasters and water safety, showing that waterborne diseases do not necessarily increase after natural disasters (Wang et al. 2009; Waring et al. 2005).

By comparing the change between IFLS3 and IFLS5, I identified the effect of the earthquake on waterborne diseases eight years after the earthquake. The result indicates that the beneficial effect on skin infection and eye infection is sustained in the long term. In the long term, the affected individual also experienced a significant reduction in vomiting, counting for a 5.17 percentage points reduction and a 17% relative decrease in the group average. The long-term impact of skin infection is even more significant than the short-term impact, giving a 12.1 percentage points reduction. The beneficial effect on eye infection continues, but the magnitude decreases in the long term. The waterborne diseases index and Romano-Wolf P-values indicate that the result is not likely random. I find that waterborne disease symptoms significantly reduced in the affected area one year after and eight years after the earthquake.

Furthermore, I identify the effect of the earthquake on non-waterborne disease symptoms. Table 4 shows that the non-waterborne disease index has a marginal reduction in the short term and decreases significantly in the long term. Specifically, no symptoms significantly change in the short term, even if they all move downward. Romano-Wolf P-value rejects all hypotheses and concludes no significant effects. Additionally, as headaches and fever reduced significantly in the long term, the non-waterborne disease index decreased in the long term. Romano-Wolf P-value provides a more conservative estimation where only headache has decreased significantly in the long term. Headache and fever could also be symptoms of water-related infection. Thus, the results indicate that none-waterborne disease symptoms do not have substantial change as the waterborne disease symptoms shown in table 3.

The results suggest a beneficial and lasting effect on the waterborne disease after the earthquake. Furthermore, although the earthquake did not affect the non-waterborne disease incidence in the short term, a significant positive effect was observed in the long term, especially for headaches and fever.

Reported Health Status of Never-movers

While natural disasters could significantly worsen the living environment and reduce the health status of affected individuals, I find the earthquake lowered waterborne diseases in the earthquake-affected communities. Would reducing the incidence of diseases improve the reported health status, or would the negative impacts of the earthquake outweigh the beneficial effect of disease incidence reduction? To answer this question, I explore how the earthquake affects the reported health status among all individuals who lived in the same community after the earthquake. Similarly, as the incidences of disease symptoms, I replaced Y_{ics} with individual reported health status from 1 to 4, in which 4 indicates the best health condition and 1 demonstrates the worst health condition.

Table 5 shows the estimation results for the short-run and long-run effects after the earthquake using OLS and Order Logit Regression. By estimating the reported health status change from IFLS3 to IFLS4, the reported health condition increases by 0.08 units counting for a 4% increase in the linear scale. The OLS result is consistent with the ordered logit regression result. This result shows a negligible but significant positive effect on health conditions one year after the earthquake or during the reconstruction project. However, when measuring the change from IFLS3 to IFLS5, not only has the magnitude of health benefit diminished, but also it became statistically insignificant. Thus, the reconstruction seems to help improve the health condition immediately after the earthquake, but this effect diminishes to a negligible level eight years after the earthquake. Based on this result, no significant adverse effects from the earthquake on the reported health have been found, implying no significant health issues among those who lived in the same affected communities.

Toothache

Another reported acute symptom is toothache. Since the incidence of toothache could either be induced by oral infection or tooth cavity, it is unclear if water safety would result in the incidence of toothache. Thus, I exclude it from the main analysis. Following the same specification as other disease symptoms, the result in table 6 indicates that toothache decreased significantly one year and eight years after the earthquake, with a relative reduction of 95% and 459% from the sample average. I did not find any credible data or resources on why toothache has reduced after the earthquake. Education on oral health from JRF and improved water quality from water facilities could be the reason, but untestable by this study (Java Reconstruction Fund, 2012).

Accessibility to A Safe Environment of Never-movers

To explain why the waterborne disease symptoms decreased significantly one year and eight years after the earthquake, I verified the change of the reported household living environment factors in earthquake-affected households and not affected households among never-movers. Thus, I studied if the accessibility to safe environment factors changed by comparing the affected and non-affected households before and after the earthquake.

Table 7 presents the estimation result. The affected households had a much higher likelihood of access to safe drinking water, safe using water, and sewage, which was the focus of the Java Reconstruction Fund (JRF) and Community Settlement Reconstruction and Rehabilitation Project (CSRRP). Among households affected by the earthquake, the likelihood of accessing safe drinking water increased by about 20.5 percentage points, 33% above the sample average. The likelihood of accessing safe using water increased by about 11.0 percentage points, counting 17% above the sample average. It is 17.7 percentage points and 28% above average for sewage service. The environment index and Romano-Wolf P-values indicate that the change is not likely random.

In the long-term, the beneficial effect of sewage diminishes to an insignificant level, but access to safe drinking and using water still generates a long-term beneficial effect. Noticeably, household access to safe drinking water increased significantly in both the short and long-term, and the effect was much more pronounced in the long term. As the Romano-Wolf multiple hypothesis testing provides conservative testing on various outcomes, safe drinking and using water significantly improved after the earthquake. Based on the result, I can conclude that the living environment as a whole improved for the affected households both in the short term and long term. Thus, this result provides evidence on why waterborne diseases decreased both in the short and long term as individuals had better access to safe drinking and using water, reducing the probability of contracting waterborne diseases by drinking and using from centralized service water than from natural water resources.

Migration and Ever-mover

Until now, I have only considered individuals and households living in the same communities over time. However, households and individuals could decide to move after the onset of natural disasters (Blanco,2023). This migration decision could affect their incidences of contracting diseases and impact the living environment. This section studies the migration decision and ever-movers in the sample to better understand if individuals decided to move after the earthquake, which affects their health conditions and living environment.

After natural disasters, migration and relocation programs are common. Reconstruction or natural disaster mitigation programs move the affected individuals to a safer location for temporary living, preventing the disease outbreak by avoiding exposure to potential risks (Jafari et al., 2011). However, JRF employed a community-centered program, providing temporary houses to affected individuals near their original residential location. In this case, relocation or temporary housing should not have involved their migration decision or forced the households to migrate to a new place after the earthquake. I followed Brown and Velásquez (2017) and constructed a household fixed effect model of studying migration decisions one year and eight years after the earthquake. IFLS reports household migration history between waves at the village, subdistrict, district, and province levels. As the earthquake is recorded at the community level, I count moving only at the village level as still living in the same community but when moving at the subdistrict level. Moreover, no information records if the household moved after the earthquake or before the earthquake between IFLS3 and IFLS4, and no data provide information on their moving destination. For simplicity, I assume that all migration between IFLS3 and IFLS4 happened after the earthquake. I explore the migration decision based on the earthquake by tracking the household over time from IFLS3 to IFLS4 and IFLS3 to IFLS5. The estimation specification is demonstrated as follows:

$$Y_{hcs} = \beta_0 + \beta_1 M_{hcs} + \theta_h + \vartheta_s + \varepsilon_{isc}$$

If a household moved to a new community after IFLS3, I count them as ever-mover in IFLS4, causing Y_{hcs} to be equal to 1 in both IFLS4 for each household y, living in community c, on survey wave s. If they did not move between IFLS3 and IFLS4 but moved between IFLS4 and IFLS5, Y_{hcs} equals 0 in IFLS4 and 1 in IFLS5. I dropped those moved between IFLS3 and IFLS4 from the sample when studying migration decisions between IFLS3 and IFLS5 as they have already moved between IFLS3 and IFLS4. M_{hcs} is the time-invariant earthquake indicator following the previous estimation. θ_h is the household fixed effect, ϑs is the survey wave fixed effect, and ε_{isc} is the error term.

Table 8 summarizes the results of the migration decision of the households. The results show that, between IFLS3 and IFLS4, the likelihood of households moving to another community increases by 9.14 percentage points, counting about a 140% increase in the sample average compared to other households not affected by the earthquake. However, the result for IFLS3 and IFLS5 is different. Those who did not move immediately after the earthquake are very unlikely to move to other communities eight years after the earthquake. Thus, exposure to an earthquake failed to predict the probability of the household moving eight years after the earthquake. This result shows that the earthquake might have forced the households to move even if JRF employs a community-centered program to help household living in the same location.

Disease Incidence and Living Environment of Ever-Movers

Given that being exposed to the earthquake predicts a higher probability of moving, would those individuals or households who moved have different conditions in disease incidences and have different living environments compared to those who have never moved? Table 9 summarizes the health status, waterborne disease index, and non-waterborne disease index changes after the earthquake for ever-movers. Unlike never-movers, ever-movers experienced worse health conditions both in the short-term and long term. The waterborne disease and non-waterborne disease indices did not change compared to those who were never exposed to the earthquake. In the long term, the non-waterborne disease index increases marginally in the short term, proposing why the health condition has worsened among ever-movers. Thus, households who decide to move might suffer from a higher likelihood of contracting infection and face a more considerable health status drop. Table 10 shows that ever-mover households have better access to sewage in the short term but worse access to other service types. In the long term, the worsening condition improves, but they still have worse access to inside-house water, safe using water, sewage, and garbage services. The total environment index indicates the worsening condition in the short term and insignificant condition in the long term. Therefore, migration might not result in a positive impact on those who moved after the earthquake.

Threats to Identification Assumption

The Inferential assumption of this study is that trends in the disease symptoms across individuals and accessibility to the safe environment across households would have been the same in communities exposed to the earthquake and communities not exposed to the earthquake had the earthquake not happened. Thus, the primary threat to my empirical strategy providing causal identification is that individuals and households had already experienced different level trends or levels in disease incidences and accessibility to safe living environments before the earthquake. To verify if the effect being estimated in the main analysis is biased by unobserved linear trends correlated with the earthquake's intensity, I estimate the same estimation of the main analysis by using data from one wave before the earthquake. I followed the same procedure but replaced the comparison wave with IFLS2 from 1997. By comparing changes between 1997 (IFLS2) and 2000 (IFLS3) and assigning the earthquake intensity levels to these observations in the same communities, I can explore if those changes or different trends had already happened before the earthquake. If the changes in disease incidence and accessibility to a safe environment were not a result of underlying linear trends with earthquake intensity, I should not observe a significant effect, as the earthquake intensity should not predict an apparent change from IFLS2 to IFLS3.

Tables 11 and 12 summarize the placebo test analysis. Firstly, the earthquake predicts a large and precise notable change in the waterborne disease symptom index, and stomachache, vomiting, skin infection, and eye infection have a higher incidence level in the earthquake-affected communities. The same effect is observed in the non-waterborne disease symptoms, including coughing, difficulty breathing, and fever. As for the environmental factors, other than better access to safe swage before the earthquake, households did not have any significant changes in terms of safe drinking and using water, water inside of the house, and access to toilet and garbage services are not statistically different. This result shows that individuals and households in earthquake-affected communities had worse disease conditions and no different living environment. Thus, this placebo test strengthens the result from my main analysis and indicates the actual effect of the reconstruction might be even higher than the current observed level.

Robustness Check

Choices of Diseases Definition

I define waterborne diseases as diseases that transmit through water, and I exclude headaches and fever as part of the waterborne disease symptoms as any infection could induce them. As part of the robustness check, I include headache and fever as the symptoms of waterborne diseases and define cough, running nose, and difficulty breathing as actual non-waterborne diseases. Table 13 summarizes the results based on the new definition. The new definition shows that exposure to the earthquake leads to a significant drop in waterborne disease symptoms one year and eight years after the earthquake, but no difference in none-waterborne diseases. The earthquake also had a beneficial effect on non-waterborne diseases eight years after the earthquake. Therefore, including fever and headache as waterborne diseases or non-waterborne diseases does not change the result.

Definition of Exposed to Earthquake

I define being exposed to the earthquake as individuals, households, and communities experiencing MMI higher than 6 as buildings and infrastructure suffer from non-negligible damage at this intensity level (Wald et al. 1999). I change the definition of experience the earthquake to MMI equal to or greater than 5 and 5.5 to include less intensity area to be defined exposed to the earthquake. Table 14 shows the results. When I define earthquake with MMI ≥ 5, the waterborne diseases index both decreased, but only significant in the long term. When I define earthquake with MMI ≥ 5.5, the waterborne diseases index drops in the short and long term, and both indices are estimated precisely. This robustness check indicates that the choice of defining experiencing the earthquake using different MMI does not affect the results.

Cost-Effectiveness Analysis (CEA)

To better understand the overall reconstruction program, this section examines the cost-effectiveness of the Java Reconstruction Fund (JRF)'s effort on water safety. To estimate the effectiveness of the program, two comparators are proposed: Java Reconstruction Fund vs. no intervention when worsening disease condition without reconstruction is not considered, and Java Reconstruction Fund vs. no intervention but worsening disease condition without reconstruction is considered. Specifically, as I cannot observe the disease condition if the reconstruction program had not happened in the earthquake-affected area, I can only assume how much the JRF has helped improve the disease condition. In this case, I assumed two conditions could have been true. First, I assumed that had the earthquake not happened, the earthquake-affected area would have had the same condition as individuals living in the control group communities. Second, I assumed that had the earthquake not happened, the earthquake-affected area would suffer from disease outbreaks over the next 20 years to count for the extreme case.[9] I use this extreme case to count for the possibility that the reconstruction and records of the best potential health benefit of the program could have averted disease outbreaks. For the first condition, I use data from the control group in IFLS4 to reflect the state of the disease. For the second condition, I followed the current literature studying typhoid and paratyphoid fever in Jakarta, Indonesia, in 2004, which indicates the odd ratio (OR) of having typhoid fever during the flood is 4.52 with 95% of CI (1.90-10.73) (Vollaard et al., 2004).

[9] 20 years is chosen to count for the extreme condition where disease outbreaks lasted the whole study period. This is a very unlikely event, as Watson et al. (2007) indicate that the disease stage could last 1-3 years, depending on how effective the reconstruction programs are.

Another challenge of CEA is that I have a vector of disease symptoms, not one disease. For simplicity, I aggregate all disease symptoms together and call this stage the disease stage. I define the disease parameter by finding the group average and standard deviation of none affected area individuals and call this parameter no intervention disease incidence. This process gives the control group average of 0.1011 and a standard deviation of 0.1687. Given that the waterborne diseases index decreased by -0.129 levels of standard deviation from the mean in the short term and -0.165 levels of standard deviation from the mean in the long term. This gives the disease stage probability of 0.07933 in the short term and 0.0732 in the long term. As there is little difference, I assume the short-term effect lasts to the long-term. As the odd ratio (OR) of having typhoid fever during the flood is 4.52, I assume that the disease stage probability is 0.4985 had no reconstruction not happened (Vollaard et al., 2004). Gramma distribution is used to count for non-negative probability.

I further assumed that the estimated reduction in quality-adjusted life year (QALY) of this disease stage is similar to other acute infections: COVID-19 with QALY reduction of 0.061 (95% CI: 0.016-0.129) (Basu & Gandhay, 2021) and Clostridium difficile infection with QALY reduction 0.050 (95% CI: 0.015-0.085) (Barbut et al., 2019). I assumed the QALY gained from averting the disease stage is 0.05 with a standard error of 0.015. I use 20 years as the study span, each year an independent event, a 3% discount rate, half-year correction, and healthy status with a QALY of 1.

The cost of the JRF is based on the Java Reconstruction Fund Progress report. With CSRRP investing USD 75.12 million for all earthquake reconstruction-related projects and the affected population is 1.5 million, the per capita cost of the program is about 30 dollars. This number will likely overestimate the cost as USD 75.12 million was invested in restoring the affected area, including housing, road, and other infrastructures. I assume the cost is a normal distribution with 30 dollars and 10 dollars standard deviation to allow an uncertain change in cost. All related parameters used for this exercise are summarized in table 15.

Figure 4 shows the project effectiveness distribution between the three projects. When the Java reconstruction program is implemented, it generates 15.04 units of QALY which is higher than no intervention situations. Figures 5 and 6 show that the project's Willingness To Pay (WTP) is 1,800 dollars, indicating the Incremental Cost-Effectiveness Ratio (ICER) between the Java reconstruction program and no intervention is 1,800 USD/QALY or affected individuals are willing to pay 1800 USD for the Java reconstruction program to be accepted. Considering Indonesia's GDP per capita was 1,572 USD in 2006, the project counts for 114.5% of the GDP per capita, which satisfies the cost-effectiveness criteria of being an effective program following the definition given by WHO (World Bank, 2023; World Health Organization, 2003). However, it is still less effective than vaccines and other small-scale programs targeting diarrhea and other waterborne diseases (Rautenberg et al., 2022; Rheingans, 2014). In addition, considering that a significant disease outbreak would have happened if no reconstruction program had been implemented, the WTP decreased to 0.3 USD/QALY, indicating JRF is a highly effective program. However, it is unlikely that such outbreaks would last for 20 years, and 0.3 USD/QALY is likely to overstate the effect largely. Nevertheless, the result still provides evidence that the actual impact of the JRF would still be lower than 1,800 USD/QALY, which might satisfy the WHO definition of highly effective program criteria with cost/QALY lower than annual GDP per capita. Combining both results, the CEA indicates that JRF provided an effective after-natural disaster strategy to manage waterborne diseases.

Discussion

Based on the result, I find evidence that the earthquake significantly reduced the incidences of waterborne disease, access to safe and used drinking water increased, and the living environment improved both in the short and long term. Those results lead to a different conclusion than other literature on natural disasters and water safety. One explanation is the successful and robust JFR implemented right after the earthquake. Unlike most other natural disasters in developing countries, the 2006 Yogyakarta earthquake had one of the most generous international assistance. On top of that, the local government organized fast and efficient after-disaster assistance and reconstruction programs in the affected area (Java Reconstruction Fund, 2007). One of the most critical responsibilities of the reconstruction program was to restore safe water and provide a secure sewage system to decrease the spreading of waterborne diseases. For this earthquake, the estimated direct damage is about USD3.1 billion dollars, and 94.06 million USD was distributed to the affected area for facilities reconstruction financial assistance (Java Reconstruction Fund, 2012). The compensation was about 30% of the total damage, about 20% more than other average relief and reconstruction programs in developing countries (Freeman et al. 2002). One of the reconstruction priorities was to repair the water supply and sewage system, provide safe water, and improve the living environment after the earthquake. Therefore, the successful reconstruction program could explain why waterborne diseases and the living environment improved after the earthquake. The Cost-Effectiveness Analysis (CEA) indicates that the WTP of the reconstruction program is about 1,800 dollars/QALY when worsening condition without reconstruction is not considered, counting for 114.5% of the GDP per capita and qualified as a cost-effective program following the definition of WHO. Moreover, this number decreases to 0.5 dollars/QALY if I consider potential diseases outbreak had the reconstruction program not been implemented. The analysis indicates that the reconstruction program is effective compared to other large-scale projects aiming to improve water safety.

Although the symptoms of waterborne disease decrease significantly both in the short-term and long-term and symptoms of non-waterborne diseases decrease in the long term, those beneficial effects of acute infection reduction are not enough to compensate for the negative impact of the earthquake. Furthermore, the individual health status did not change for those who did not move or move eight years after the earthquake. Therefore, my results show that this program is not enough to promote better health conditions in the affected area long-term, even if the reconstruction project helped to lower acute infections. Programs focusing on other aspects of the health conditions, such as reducing chronic diseases and providing medical services, are needed to improve the health status in the affected area in the long term. Therefore, more effectively designed reconstruction programs focusing on long-term and non-water-related programs after natural disasters might still be beneficial.

One of the most significant weaknesses of this study is the lack of actual earthquake damage data at specific geographic locations. Although MMI is a standard indicator for earthquake damage, it does not accurately reflect the earthquake's effect on individuals and households. With the development of remote sensing technology, aerial and satellite images of the affected areas might provide more accurate and direct information by detecting before and after the earthquake change (Fan et al., 2019). Therefore, using datasets like Landsat and other satellite images could help more accurately identify the study area changes. However, as the correct-line sensor failed for Landsat during the earthquake period and the expensive cost of private satellite images, this study cannot implement those techniques to measure the earthquake accurately. Additionally, no spatial data records the distribution of the reconstruction funds and resources, which does not allow me to study the actual effect of the reconstruction. Therefore, it calls for more advanced studies on this topic and explores the impact of rebuilding after natural disasters.

Conclusion

What happens to water safety after major earthquakes, and would earthquakes result in more waterborne diseases? Could reconstruction-induced improvement in the household living environment compensate for the negative impacts of earthquakes? To provide an answer to those questions. I use the 2006 Yogyakarta Earthquake and its reconstruction program as a natural experiment to identify the effect of natural disasters and their reconstruction on water safety. Using data from different data sources and implementing a different-in-different study design, I found that waterborne disease symptoms decreased significantly one year and eight years after the earthquake and the affected household had access to more safe drinking and using water one year and eight years after the earthquake.

Moreover, by comparing movers and non-movers after the earthquake, although moved households access to the sewage system temporarily, they still faced a worse living environment, higher disease incidence, and a higher reduction in health status right after the earthquake. This result also indirectly shows that the reconstruction program might have effectively improved the water supply system in the earthquake-affected area since the welfare related to safe water did not increase for those who moved out of the original communities. Thus, if the reconstruction is robust, moving after natural disasters might not be a good choice for the affected individuals and households.

Lastly, as there is no detailed spatial distribution of information on the assistance program, I cannot identify the causal relationship between earthquake reconstruction on disease and health conditions among affected individuals. However, the effect of earthquake reconstruction on water facility repair/rebuild to the local communities could potentially explain my results. This research complements the literature on natural disasters, water safety, and reconstruction. It explores the long-term effect of natural disasters on water safety and proposes that a more advanced study on an extensive reconstruction program on water safety is needed.

Tables & Figures

Table 1: Communal Difference Between Study and Excluded Area

	IFLS3		IFLS4	
	Study Area Communities	Excluded Communities	Study Area Communities	Excluded Communities
Number of Health Posts	9.434	6.067	10.145	6.311
Annual Health Stuff Visit	70.224	45.915	34.832	26.911
=1 If this community has slums	0.264	0.158	0.308	0.344
=1 If this community has smelly air	0.244	0.027	0.289	0.063
=1 If this community has exposed garbage	0.153	0.190	0.231	0.166
=1 If this community has exposed manure	0.061	0.209	0.100	0.181
=1 If this community has blocked water	0.258	0.291	0.269	0.299
=1 If this community has still water	0.120	0.300	0.147	0.370
=1 If this community has roaming cattle	0.854	0.719	0.828	0.728
=1 If this community cleans the yard	0.410	0.583	0.368	0.522
=1 If this community cares for grass	0.852	0.812	0.802	0.836
=1 If this community has flies near food	0.148	0.179	0.198	0.164

Note: The number of health posts indicates how many health posts are within the community and the annual health staff visit indicates how many health personnel travels to those posts per year. The rest of the nine community environment variables were observed by the interviewers during the survey.

Table 2: Summary Statistics

	IFLS3			IFLS4			IFLS5		
Variables	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
				Individual Level					
age	9,975	26.961	18.782	9,138	34.982	19.045	6,703	41.992	18.245
MMI	9,975	2.800	0.000	9,138	3.627	1.321	7,051	3.668	1.358
Stomachache	9,969	0.222	0.415	9,115	0.210	0.407	6,693	0.296	0.457
Vomiting	9,968	0.104	0.305	9,116	0.100	0.300	6,694	0.139	0.346
Diarrhea	9,969	0.099	0.299	9,116	0.066	0.249	6,691	0.097	0.296
Skin Infection	9,969	0.117	0.321	9,116	0.083	0.276	6,695	0.146	0.353
Eye Infection	9,969	0.057	0.232	9,116	0.046	0.210	6,695	0.065	0.246
Headache	9,969	0.509	0.500	9,115	0.503	0.500	6,690	0.581	0.493
Running Nose	9,969	0.530	0.499	9,116	0.398	0.489	6,693	0.441	0.497
Cough	9,968	0.402	0.490	9,116	0.324	0.468	6,695	0.401	0.490
Difficult Breathing	9,969	0.067	0.250	9,116	0.066	0.249	6,695	0.086	0.280
				Household Level					
Safe Drinking Water	2,657	0.609	0.488	2,598	0.629	0.483	2,486	0.586	0.493
Water inside of House	2,662	0.433	0.496	2,625	0.532	0.499	2,554	0.617	0.486
Safe Using Water	2,657	0.577	0.494	2,598	0.703	0.457	2,486	0.798	0.402
Access to Toilet	2,657	0.769	0.422	2,598	0.859	0.348	2,486	0.941	0.236
Have Sewage	2,657	0.578	0.494	2,598	0.649	0.477	2,486	0.601	0.490
Garbage Service	2,662	0.301	0.459	2,625	0.341	0.474	2,554	0.383	0.486

Table 3: Incidence of the Waterborne Disease Symptoms among Never-movers

Short-term Effect of the Earthquake on Waterborne Diseases (IFLS3 to IFLS4)

	Waterborne Disease Index	Stomachache	Vomiting	Diarrhea	Skin Infection	Eye Infection
Earthquake (MMI≥6)	-0.129***	-0.0235	-0.0154	-0.0240	-0.0845***	-0.0498***
	(0.0354)	(0.0212)	(0.0198)	(0.0170)	(0.0229)	(0.0163)
Observations	14,262	14,263	14,262	14,263	14,263	14,263
R-squared	0.012	0.011	0.004	0.015	0.013	0.008
Number of Individuals	7,143	7,143	7,143	7,143	7,143	7,143
Age Control	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES
Mean of Y	-0.00015	0.209	0.0998	0.0825	0.104	0.0538
Romano-Wolf P-value	N/A	0.3564	0.3564	0.2277	0.0099	0.0099

Long-term Effects of the Earthquake on Waterborne Diseases (IFLS3 to IFLS5)

	Waterborne Disease Index	Stomachache	Vomiting	Diarrhea	Skin Infection	Eye Infection
Earthquake (MMI≥6)	-0.165***	-0.0334	-0.0517***	-0.0169	-0.121***	-0.0358**
	(0.0304)	(0.0313)	(0.0193)	(0.0221)	(0.0246)	(0.0168)
Observations	9,394	9,397	9,397	9,398	9,395	9,398
R-squared	0.013	0.043	0.043	0.012	0.011	0.016
Number of Individuals	4,834	4,834	4,834	4,834	4,834	4,834
Age Control	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES
Mean of Y	-0.0114	0.240	0.120	0.0969	0.134	0.0605
Romano-Wolf P-value	N/A	0.3663	0.0198	0.3663	0.0099	0.0396

Robust standard errors are clustered at the community level. All estimations include age controls, the interview month fixed effect, the individual fixed effect, and the survey wave fixed effect. Age controls include age and age square to record linear and nonlinear effects on disease incidence from age. Survey Wave FE controls the difference in response to the incidence of the diseases across waves. Interview Month FE and Individual FE capture the unobservable characteristics that would affect the incidence of the symptoms due to the monthly and individual characteristics. Roman-Wolf P-value is based on Romano-Wolf Step-Down adjusted P-Values calculation.

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Incidence of the Nonwaterborne Disease Symptoms among Never-movers

Short-term Effect of the Earthquake on Non-Water Related Diseases (IFLS3 to IFLS4)

	Non-waterborne Disease Index	Headache	Running Nose	Cough	Diffcult Breathing	Fever
Earthquake (MMI≥6)	-0.0837*	-0.0250	-0.0405	-0.0540	-0.00404	-0.0260
	(0.0460)	(0.0348)	(0.0507)	(0.0425)	(0.0135)	(0.0206)
Observations	14,263	14,263	14,263	14,263	14,263	14,263
R-squared	0.008	0.005	0.041	0.014	0.002	0.028
Number of Individuals	7,143	7,143	7,143	7,143	7,143	7,143
Age Controls	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES
Mean of Y	-0.00136	0.503	0.461	0.367	0.0669	0.217
Romano-Wolf P-value	N/A	0.5446	0.5446	0.3663	0.6832	0.3663

Long-term Effect of the Earthquake on Non-Water Related Diseases (IFLS3 to IFLS5)

	Non-waterborne Disease Index	Headache	Running Nose	Cough	Diffcult Breathing	Fever
Earthquake (MMI≥6)	-0.144***	-0.125***	-0.0430	-0.0393	-0.00922	-0.0519*
	(0.0381)	(0.0330)	(0.0401)	(0.0398)	(0.0136)	(0.0291)
Observations	9,391	9,393	9,396	9,398	9,398	9,397
R-squared	0.007	0.033	0.026	0.008	0.008	0.026
Number of Individuals	4,834	4,834	4,834	4,834	4,834	4,834
Age Controls	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES
Mean of Y	-0.0166	0.539	0.480	0.398	0.0765	0.225
Romano-Wolf P-value	N/A	0.0099	0.3762	0.3762	0.4455	0.1188

Robust standard errors are clustered at the community level. All estimations include age controls, the interview month fixed effect, the indivudal fixed effect, and the survey wave fixed effect. Age controls include age and age square to record linear and nonlinear effect on dieases incidenes from age. Survey Wave FE controls the difference in response to the incidence of the diseases across waves. Interview Month FE and Individual FE capture the unobservable characters that would affect the incidence of the symptoms due to the monthly and individual characteristics. Roman-Wolf P-value is based on Romano-Wolf Step-Down adjusted P-Values calculation, which provides a conservetive estimiation on hypotheiss testing of multiple outcomes.

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Reported Health Status Change Among Never Movers

	IFLS3 to IFLS4		IFLS3 to IFLS5	
Reported Health Condition	OLS	Ordered Logit	OLS	Ordered Logit
Earthquake (MMI≥6)	0.0818**	0.541***	0.0218	0.0633
	(0.0317)	(0.202)	(0.0368)	(0.186)
Observations	14,264	5,182	9,399	4,394
R-squared	0.009	0.0195	0.035	0.0596
Number of Individuals	7,143	2444	4,834	1995
Age Controls	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES
Mean of Y	2.034	2.044	2.067	2.110

Robust standard errors are clustered at the community level. The reported health condition has four levels: 1 is very bad, 2 is bad, 3 is healthy, and 4 is very healthy. Interview Age controls include age and age square. Survey Wave FE controls the difference in response to the incidence of the diseases across waves. Interview Month FE and Individual FE capture the unobservable characters that would affect the incidence of the symptoms due to the monthly and individual characters. Ordered logit regression utilizes the log conditional likelihood estimation method, which only counts groups with a variation on the dependent variable. Therefore, the number of observations is much smaller when ordered logit regression is applied.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: The Effect of Earthquake on Toothache

Toothache	From IFLS3 to IFLS4	From IFLS3 to IFLS5
Earthquake (MMI≥6)	-0.0603**	-0.0979***
	(0.0233)	(0.0342)
Observations	14,263	9,398
R-squared	0.004	0.005
Number of Individuals	7,143	4,834
Age Controls	YES	YES
Fixed Effects	YES	YES
Mean of Y	0.0646	0.0213

Note: Robust standard errors are clustered at the community level. All estimations include age controls, the interview month fixed effect, the individual fixed effect, and the survey wave fixed effect. Toothache is defined as experiencing any toothache over the past four weeks.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Accessibility to Safe Environment Among Never-movers

Short-term Effect of The Earthquake on Access to Safe Environment (IFLS3 to IFLS4)

	Environment Index	Safe Drinking Water	Water Inside House	Safe Use Water	Access to Toilet	Have Sewage	Garbage Service
Earthquake (MMI≥6)	0.175**	0.205***	-0.0635	0.110**	-0.00606	0.177***	0.0690
	(0.0709)	(0.0570)	(0.0476)	(0.0499)	(0.0275)	(0.0506)	(0.0658)
Observations	4,627	4,627	4,652	4,627	4,627	4,627	4,652
R-squared	0.013	0.016	0.030	0.075	0.049	0.023	0.025
Number of Households	2,324	2,324	2,326	2,324	2,324	2,324	2,326
Fixed Effects	YES	YES	YES	YES	YES	YES	YES
Mean of Y	0.0194	0.620	0.484	0.645	0.824	0.614	0.333
Romano-Wolf P-value	N/A	0.0099	0.2475	0.0396	0.7723	0.0099	0.2673

Long-term Effect of The Earthquake on Access to Safe Environment (IFLS3 to IFLS5)

	Environment Index	Safe Drinking Water	Water Inside House	Safe Use Water	Access to Toilet	Have Sewage	Garbage Service
Earthquake (MMI≥6)	0.154*	0.277***	-0.0418	0.107**	0.000101	-0.0167	0.103
	(0.0793)	(0.0632)	(0.0566)	(0.0535)	(0.0451)	(0.0373)	(0.0848)
Observations	4,391	4,391	4,448	4,391	4,391	4,391	4,448
R-squared	0.012	0.020	0.086	0.183	0.140	0.002	0.062
Number of Households	2,222	2,222	2,224	2,222	2,222	2,222	2,224
Fixed Effects	YES	YES	YES	YES	YES	YES	YES
Mean of Y	0.0182	0.601	0.533	0.689	0.861	0.584	0.349
Romano-Wolf P-value	N/A	0.0099	0.7228	0.0495	1.0000	0.8020	0.3564

Robust standard errors are clustered at the community level. All estimations include household fixed effect and survey wave fixed effect. Survey Wave FE controls to the difference in response to the environment survey across survey waves. Household FE captures the unobservable characteristics that would affect the accessibility to a safe environment. Romano-Wolf P-value provides a conservative estimation of hypothesis testing on multiple outcomes.

*** p<0.01, ** p<0.05, * p<0.1

Table 8: The Effect of Earthquake on Household Migration Decision

Migration	From IFLS3 to IFLS4	From IFLS3 to IFLS5
Earthquake (MMI≥6)	0.0914***	0.00958
	(0.0321)	(0.00619)
Observations	5,205	4,794
R-squared	0.095	0.010
Number of Households	2,662	2,551
Fixed Effects	YES	YES
Mean of Y	0.0646	0.0213

Note: Robust standard errors are clustered at the community level. If a household migrated to another community, they are counted as migrated after the earthquake. If someone migrated between IFLS3 and IFLS4, they are dropped out from the regression when estimating the migration decision between IFLS3 and IFLS5. Fixed effects include survey wave fixed effect and household fixed effect. Survey Wave FE controls the difference in migration decisions across waves. Household FE captures the time-invariant unobservable characters that would affect the migration decision.
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Health and Disease Symptoms Among Ever-movers

Short-term Effect of the Earthquake on Waterborne Diseases (IFLS3 to IFLS4)

	Health Status	Waterborne Disease Index	Nonwaterborne Disease Index
Earthquake (MMI≥6)	-1.971***	0.229	-0.0966
	(0.349)	(0.335)	(0.204)
Observations	4,087	4,086	4,087
R-squared	0.023	0.020	0.020
Number of Individuals	2,490	2,490	2,490
Age Controls	YES	YES	YES
Fixed Effects	YES	YES	YES
Mean of Y	2.007	-0.00456	0.113

Long-Run Effects of the Earthquake on Waterborne Diseases (IFLS3 to IFLS5)

	Health Status	Waterborne Disease Index	Nonwaterborne Disease Index
Earthquake (MMI≥6)	-1.666**	0.240	0.496**
	(0.734)	(0.512)	(0.231)
Observations	5,595	5,592	5,594
R-squared	0.015	0.017	0.017
Number of Individuals	3,788	3,787	3,788
Age Controls	YES	YES	YES
Fixed Effects	YES	YES	YES
Mean of Y	2.024	0.0121	0.131

Note: Robust standard errors are clustered at the community level. Induvial symptoms are not reported. Unlike never-movers, the earthquake did not result in a large change in the disease symptoms for ever-movers. The exceptions are: the incidence of diarrhea increased by 40.4 percentage points (P-value<0.01), increased eye infections by 62.2 percentage points (P-value<0.01), increased headache by 40.4 percentage points (P-value<0.1), increased cough by 70.4 percentage points (P-value<0.05), decreased difficulty breathing by 33.9 percentage points (P-value<0.05) from IFLS3 to IFLS4. The incidence of diarrhea increased by 40.4 percentage points (P-value<0.01), eye infections increased by 74.5 percentage points (P-value<0.01), increased headache by 80.7 percentage points (P-value<0.1), increased running nose by 80.7 percentage points (P-value<0.1), decreased difficulty breathing by 56.6 percentage points (P-value<0.1) from IFLS3 to IFLS5. Age controls include age and age square. Survey Wave FE controls the difference in response to the incidence of the diseases across waves. Interview Month FE and Individual FE capture the unobservable characteristics that would affect the incidence of the symptoms due to the monthly and individual characteristics.
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Accessibility to Safe Environment Among Ever-movers

Short-term Effect of The Earthquake on Access to Safe Environment

	Environment Index	Safe Drinking Water	Water Inside House	Safe Use Water	Access to Toilet	Have Sewage	Garbage Service
Earthquake (MMI≥6)	-0.202***	-0.0377	-1.116***	-0.142***	-0.156***	0.854***	-0.0602*
	(0.0462)	(0.0754)	(0.101)	(0.0395)	(0.0338)	(0.0535)	(0.0295)
Observations	429	429	434	429	429	429	434
R-squared	0.022	0.005	0.049	0.082	0.101	0.069	0.041
Number of Households	216	216	217	216	216	216	217
Fixed Effects	YES	YES	YES	YES	YES	YES	YES
Mean of Y	-0.245	0.566	0.442	0.545	0.655	0.536	0.0576
Romano-Wolf P-value	N/A	0.11880	0.00990	0.00990	0.00990	0.00990	0.00990

Long-term Effect of The Earthquake on Access to Safe Environment

	Environment Index	Safe Drinking Water	Water Inside House	Safe Use Water	Access to Toilet	Have Sewage	Garbage Service
Earthquake (MMI≥6)	-0.605	-0.213	-0.787**	-0.519*	0.0402	-0.185**	-0.137**
	(0.367)	(0.285)	(0.293)	(0.284)	(0.287)	(0.0697)	(0.0526)
Observations	487	487	498	487	487	487	498
R-squared	0.020	0.041	0.043	0.134	0.257	0.081	0.129
Number of Households	277	252	254	252	252	252	254
Fixed Effects	YES	YES	YES	YES	YES	YES	YES
Mean of Y	-0.198	0.534	0.438	0.614	0.747	0.563	0.135
Romano-Wolf P-value	N/A	0.6337	0.6040	0.6040	0.7426	0.6040	0.6040

Note: robust standard errors are clustered at the community level. Fixed effects include the survey wave fixed effect and household fixed effect. Survey Wave FE controls different responses to the environment survey across survey waves. Household FE captures the unobservable characters that would affect the accessibility to a safe environment based on different households.
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Threats to Inferential Validity for Dieases

Placebo Test on Waterborne Diseases (IFLS2 to IFLS3)

	Waterborne Disease Index	Stomachache	Vomiting	Diarrhea	Skin Infection	Eye Infection
Earthquake (MMI≥6)	0.179***	0.0965***	0.0275*	0.0324	0.106***	0.0230**
	(0.0547)	(0.0251)	(0.0163)	(0.0231)	(0.0287)	(0.0110)
Observations	17,545	17,565	17,560	17,552	17,565	17,566
R-squared	0.019	0.015	0.006	0.008	0.012	0.006
Number of Individuals	8,786	8,786	8,786	8,786	8,786	8,786
Age Controls	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES
Mean of Y	0.0325	0.210	0.102	0.0890	0.111	0.0545
Romano-Wolf P-value	N/A	0.210	0.102	0.0890	0.111	0.0545

Placebo Test on Non-waterborne Diseases (IFLS2 to IFLS3)

	Non-waterborne Disease Index	Headache	Running Nose	Cough	Difficulty Breathing	Fever
Earthquake (MMI≥6)	0.143**	0.0432	0.0415	0.0886**	0.0378**	0.0924***
	(0.0551)	(0.0363)	(0.0404)	(0.0423)	(0.0156)	(0.0239)
Observations	17,543	17,564	17,565	17,565	17,546	17,564
R-squared	0.009	0.008	0.007	0.008	0.005	0.008
Number of Individuals	8,786	8,786	8,786	8,786	8,786	8,786
Age Controls	YES	YES	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES	YES	YES
Mean of Y	-0.012	0.505	0.508	0.377	0.0726	0.248
Romano-Wolf P-value	N/A	0.8416	0.8416	0.8416	0.8416	0.8119

Note: Results follow the main result specification at the individual level.
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Threats to Inferential Validity for Living Environment

Placebo Test on the Living Environment (IFLS2 to IFLS3)							
	Environment Index	Safe Drinking Water	Water Inside House	Safe Use Water	Access to Toilet	Have Sewage	Garbage Service
Earthquake (MMI≥6)	0.0296	0.0142	-0.0471	0.0363	0.0246	0.0764*	-0.00720
	(0.0366)	(0.0423)	(0.0322)	(0.0396)	(0.0359)	(0.0446)	(0.0192)
Observations	4,968	4,969	4,979	4,968	4,969	4,969	4,979
R-squared	0.001	0.027	0.016	0.023	0.012	0.005	0.006
Number of Households	2,649	2,649	2,653	2,649	2,649	2,649	2,653
Fixed Effects	YES	YES	YES	YES	YES	YES	YES
Mean of Y	-0.00771	0.576	0.400	0.546	0.748	0.567	0.295
Romano-Wolf P-value	N/A	0.8713	0.9109	0.8713	0.8515	0.7525	0.8713

Note: Results follow the main result specification at the household level.
*** p<0.01, ** p<0.05, * p<0.1

Table 13: Different Definitions of Waterborne Disease

Short-term Effect of the Earthquake

	Waterborne Disease Index	Nonwaterborne Disease Index
Earthquake (MMI≥6)	-0.109***	-0.0697
	(0.0322)	(0.0705)
Observations	14,262	14,263
R-squared	0.012	0.007
Number of Individuals	7,143	7,143
Age Controls	YES	YES
Fixed Effects	YES	YES
Mean of Y	-0.00419	0.00184

Long-term Effects of the Earthquake

	Waterborne Disease Index	Nonwaterborne Disease Index
Earthquake (MMI≥6)	-0.173***	-0.0679
	(0.0303)	(0.0530)
Observations	9,390	9,396
R-squared	0.015	0.007
Number of Individuals	4,834	4,834
Age Controls	YES	YES
Fixed Effects	YES	YES
Mean of Y	-0.0143	-0.0101

Note: Results follow the main result specification at the individual level. Waterborne diseases now include fever and headache, as they could be caused by waterborne infection. Non-waterborne disease index includes only cough, running nose, and difficulty breathing.
*** p<0.01, ** p<0.05, * p<0.1

Table 14: Different Definitions of Earthquake using Different MMI

	Waterborne Diseases Index			
	IFLS4	IFLS5	IFLS4	IFLS5
Earthquake =1 if	MMI>=5		MMI>=5.5	
	(1)	(2)	(3)	(4)
Earthquake	-0.0589	-0.114***	-0.0959***	-0.146***
	(0.0366)	(0.0296)	(0.0364)	(0.0283)
Observations	14,262	9,394	14,262	9,394
R-squared	0.010	0.012	0.011	0.013
Number of individuals	7,143	4,834	7,143	4,834
Age Controls	YES	YES	YES	YES
Fixed Effects	YES	YES	YES	YES
Mean of Y	0.000749	-0.0139	-0.006	-0.0123

Note: Regression follows the main result at the individual level. The index control group changed based on the definition of the earthquake. Defining earthquake equals 1 if earthquake intensity is greater than 5 or 5.5. Results are robust as the main results are the exception that no observed symptoms decreased in the short term if we define earthquake as community MMI greater than 5.
*** p<0.01, ** p<0.05, * p<0.1

Table 15: Parameters Used For CEA

	Variables	**Distribution**	**Mean**	**St. Deviation**
	Prob of disease stage	Gamma	0.0758	0.1687
Java Reconstruction Fund	Cost Per Person	Normal	30.00	10.00
	Prob of disease stage	Gamma	0.1011	0.1687
No Intervention	Cost Per Person	Fixed Number	0.00	0.00
	Prob of disease stage	Gamma	0.4985	0.1687
No Intervention with Increased Diseases incidences	Cost Per Person	Fixed Number	0.00	0.00
	QALY of diseases Stage	Normal	0.95	0.015
Health Indicators	QALY of Health	Fixed Number	1.00	0.00

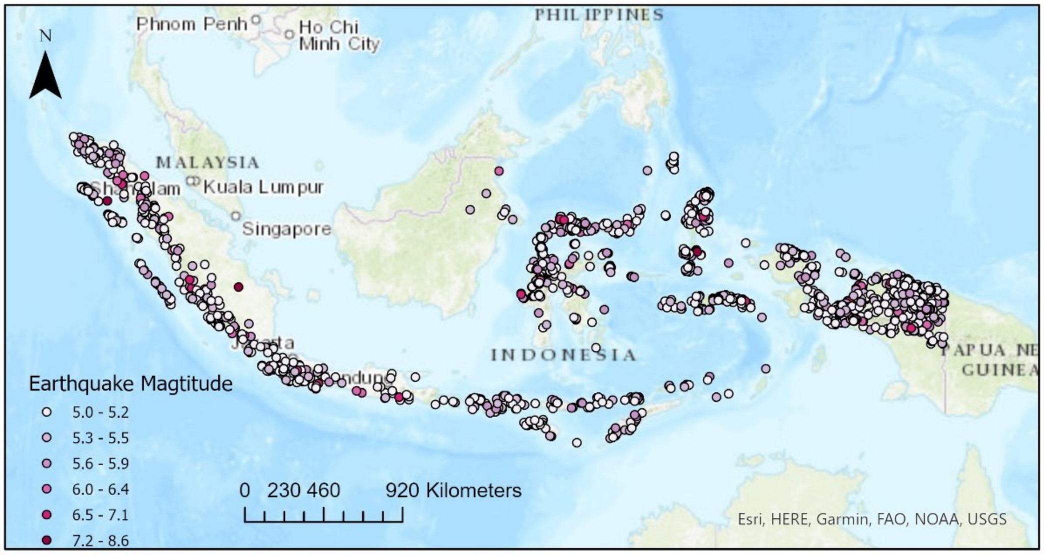

Figure 1: Onshore Earthquakes in Indonesia from 1980 to 2015

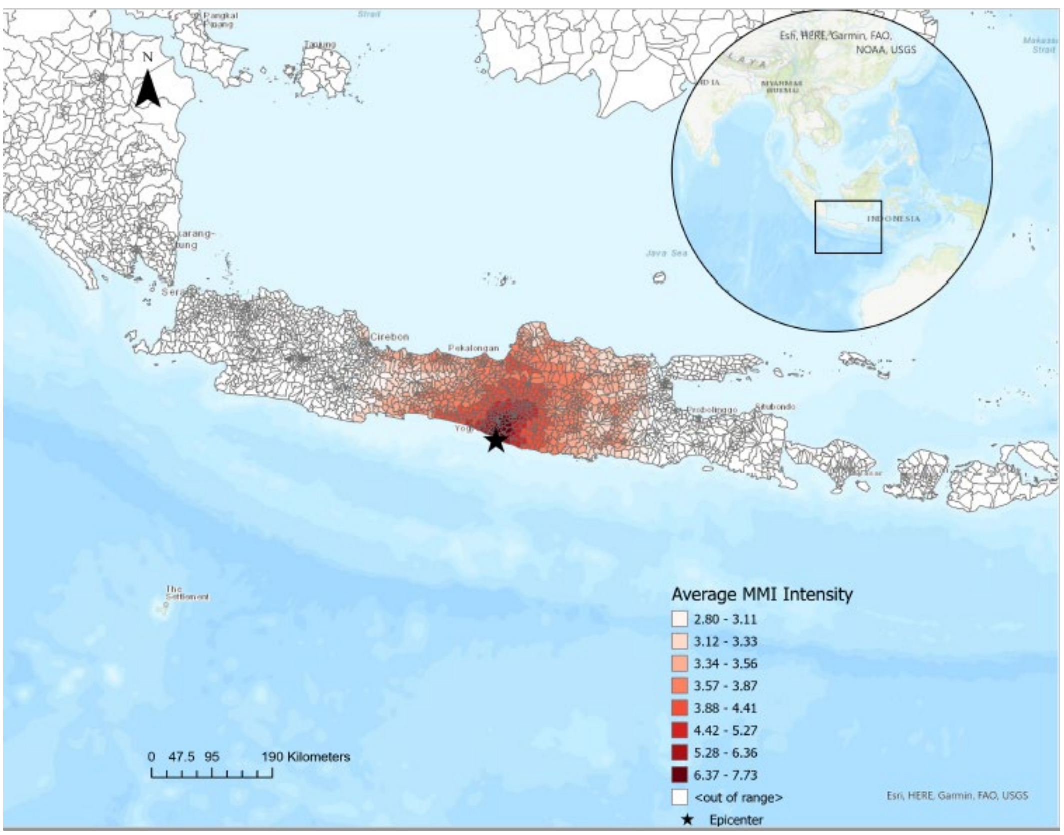

Figure 2: The 2006 Yogyakarta Earthquake Modified Mercalli Intensity

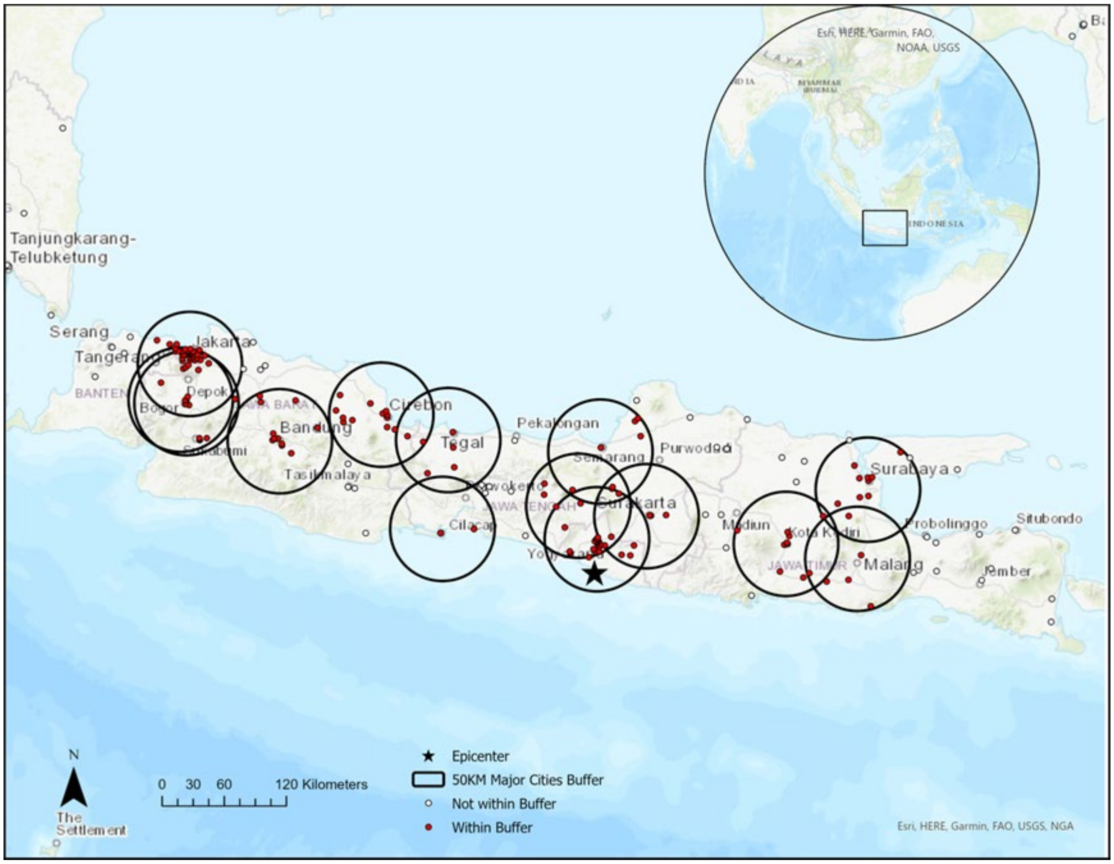

Figure 3: Communities Located Within 50 KM of Major Cities in 2000

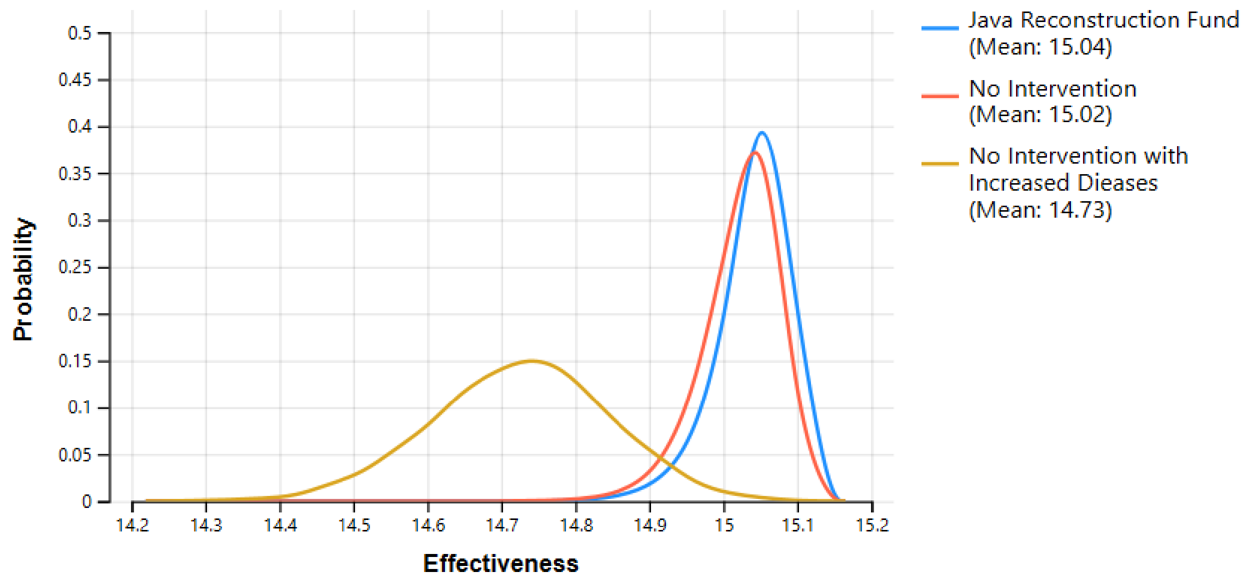

Figure 4: Project Effectiveness Distribution

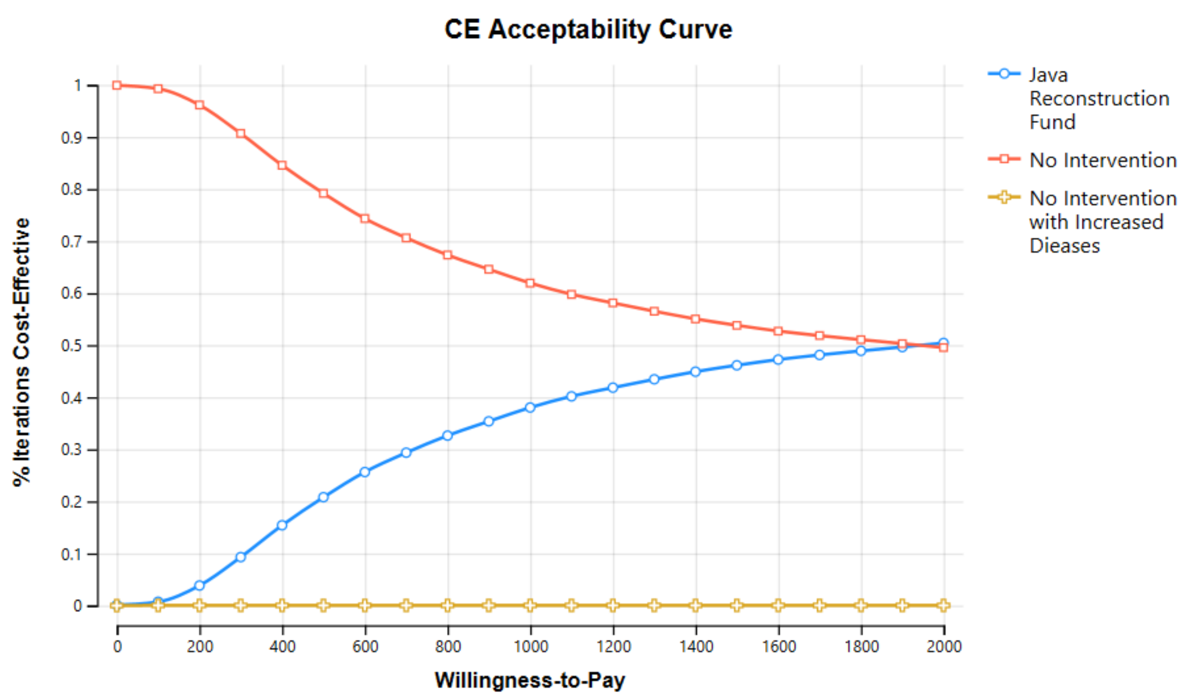

Figure 5: Program Acceptability Curve

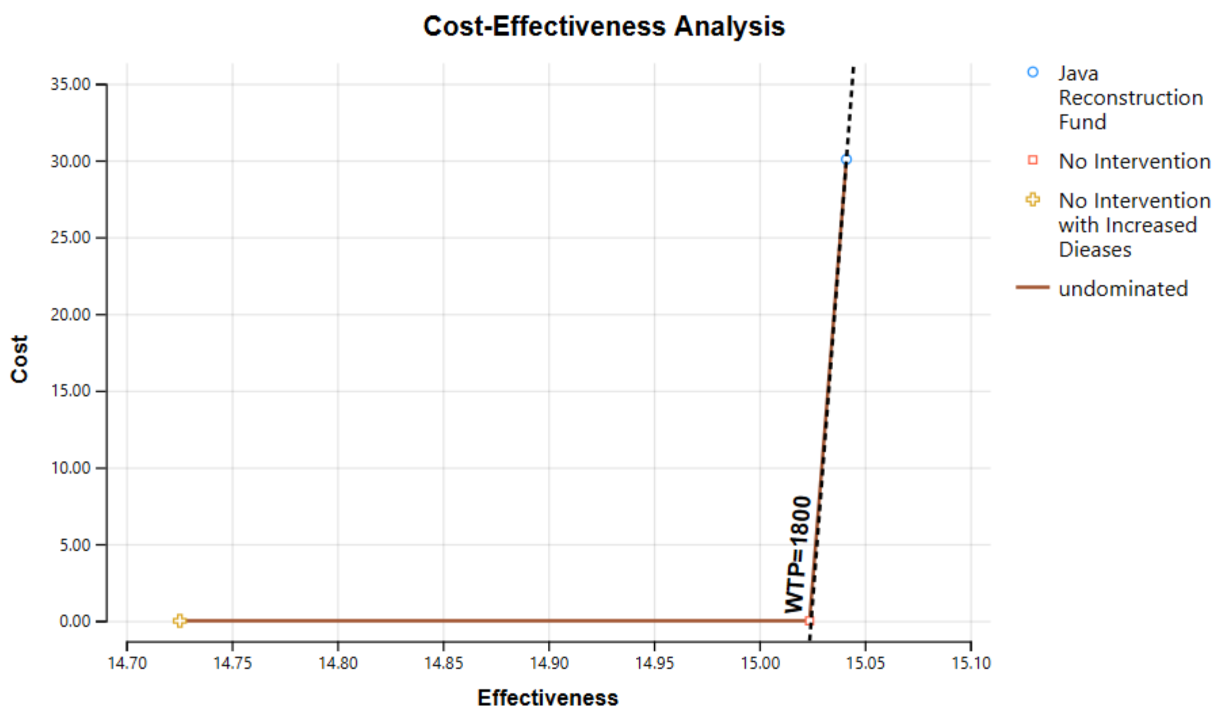

Figure 6: Willingness to Pay for the Project with Different Conditions

CHAPTER TWO: NATURAL DISASTERS AND CHILD MARRIAGE

Child marriage, defined as a first marriage before the age of 18, is a violation of human rights and has negative effects on the education and health of women and their children (Field and Ambrus,2008; Chari et al.,2017; Sekhri and Debnath,2014; Garcia-Hombrados,2021). Despite child marriage being illegal in most countries, 36 percent of women aged 20 to 24 are still married before the age of 18 in poor regions (UNICEF,2021). Understanding the factors that contribute to child marriage is crucial to designing effective policies aimed at reducing it. On top of the current grave condition on child marriage, natural disasters might have contributed to a worsening condition to child marriage (Corno, Hildebrandt, & Voena 2020, Das and Dasgupta 2020; Kumala Dewi & Dartanto, 2018). Households might marry their children at a younger age in exchange for additional income to protect children from sexual violence, due to disasters induced income shock and unsafe environment (Corno, Hildebrandt, & Voena 2020; El Arab & Sagbakken, 2019; Buchmann et al., 2021). Since 1994, more than five billion people have been affected by natural disasters, which claimed nearly two million lives and cost US$2.5 trillion in economic losses worldwide (United Nations Office for Disaster Risk Reduction, 2020). Within countries, natural disasters disproportionately affect poor people and prevent them from escaping poverty (Hallegatte et al., 2017). Natural disasters destroy physical capital, disrupt economic production and daily life, displace large groups of the population, and all those impacts could contribute to the child marriage in those regions (Bellizzi, et al., 2021). Yet, little is known about how affected individuals and households adjust decisions over the timing of children's marriage in response to natural disasters. The objective of this paper is to study how child marriage is affected by natural disasters, which are a major source of natural risk in developing countries.

To shed light on this issue, we examine the effect of earthquakes on child marriage in Indonesia, a country that is highly prone to natural disasters and has a large population of child brides. Indonesia locates in the Pacific Ring of Fire and more than 60 % of total population live in earthquake-prone areas (Suprapto, et al., 2015). As such, Indonesia is one of the world's most seismologically active countries and experience and ranks third in mortality risk from earthquakes (United Nations Office for Disaster Risk Reduction, 2020). Figure 1 illustrates the earthquakes happened in the region of Indonesia from 1980 to 2015 with magnitude higher than 5.5 Richter scale. On the other hand, 1 out of 5 women marry before the age of 18 in Indonesia, which has the eighth-largest child marriage population worldwide (UNICEF, 2019). With such high frequent earthquakes and large child marriages, Indonesia provides interesting empirical evidence on the causal relationship between earthquakes and child marriages.

To estimate the impact of earthquakes on child marriage, we combine data from various waves of a longitudinal household survey with data on earthquake intensity from the United States Geology Survey (USGS) Earthquake Catalog. First, we construct an individual-level data set with the marital and migration histories of roughly 17,000 men and women born between 1978 and 1996 by combining information from all survey waves. Next, we merged the individual-level data set with precise measurement of earthquake intensity at the subdistrict level using the USGS ShakeMaps from 1990 to 2014 to capture all earthquakes that the individual was exposed to between the age of 12 to 17. An empirical challenge is that the administrative boundaries changed over time, and we created a new measure of administrative unit that is consistent across waves to count for regional fixed effects. By tracking the individual level migration history and matching with the earthquake intensity data, we precisely measure individual exposure to earthquakes at the yearly level. Next, we leverage the panel nature of the household survey, assign before marriage household characteristics to control for omitted variable bias. As a result, we created a person-age panel data set that links the probability of getting married at a certain age with the incidence and intensity of earthquakes in the subdistrict of residence in that year.

Despite earthquakes being frequent in Indonesia, they are still rare and unpredictable to a given location, which means that households cannot anticipate them and prepare ex-ante. Our identification strategy takes advantage of the quasi-random location and timing of the earthquakes to estimate their impact on the annual hazard of child marriage (Gignoux and Menendez, 2016). We estimate the impact of experiencing an earthquake on the annual hazard of child marriage by exploiting within-age group variation in earthquakes across subdistricts as well as within-subdistrict variation in age of exposure to the earthquake. Girl who experience an earthquake of medium intensity or higher between ages 12 and 17 are 1.1 to 1.2 percentage points (about 98.8% reduction from the sample average) less likely to get married in the same year in the urban area and 5.0 to 7.3 percentage points (about 137.9% increase from the sample average) more likely to get married in the rural area. The estimated effect for boys is close to 1 percentage points but statistically insignificant at the conventional level.

Next, we explore how these effects vary with age of the child at the time of the earthquake, if the effect is instantaneous or lagged, and whether being enrolled in school at the time of the earthquakes mutes or reinforces the impact of the natural hazard on child marriage. We estimate the impact of experiencing an earthquake on the probability of getting married by age group and found that the results are driven by women who experience an earthquake between ages 12 and 15. This result indicates that in the event of earthquake, extremely young girls are those at highest risk of entering child marriage. Although age 16 is the legal minimum age for marriage for girls with parental consent in Indonesia, we find that parents marry off daughters at younger ages after experiencing an earthquake.[10] The effect is instantaneous (within the same year): experiencing an earthquake increases the probability of getting married in the same year among girls living in rural areas, with no delayed effects in the following years.[11] At the time of earthquake, girls who lived in the urban area and attended school in the previous year have 0.33 percentage points reduction in annual hazard entering the marriage in the earthquake year compared to those who were not in school in the previous year, but girls living in the rural area has a 3.4 percentage points increase in annual hazard ratio compared to those who were not in school last year. But both of those estimations are not statistically significant at the conventional level, and we cannot conclude being in school last year generates an differential effect on marriage hazard for girls at the time of earthquake compared to those who were not. Additionally, we find that the effect of the earthquake on the annual hazard into the first birth is qualitatively similar to the effect on child marriage: experiencing an earthquake between age 12 and 17 delays the first pregnancy among girls living in urban areas; experiencing an earthquake expedites the first pregnancy among girls living in rural areas, but the effect is not precisely estimated.

Corno, Hildebrandt, and Voena (2020) show that child marriage responds to aggregate economic shocks, and the sign of the response depends on the direction of marriage payments: droughts reduce child marriage in India, where dowry is prevalent, but increase child marriage in

[10] The legal minimum age for marriage without parental consent is 21 for both boys and girls. The legal minimum age for marriage for boys with parental consent is 19.
[11] The estimate on the first lag is very imprecisely estimated and not robust across different specifications.

Sub-Saharan Africa, where bride price is the dominant culture practice. We examine whether the direction of marriage payments determines how child marriage responds to earthquakes using variation in the practice of bride price across ethnic groups in Indonesia (Ashraf et al., 2019). Indonesia is a country with large group practice bride price and the bride price is larger than a one-year income of groom (Anderson, 2007). Thus, household makes decision on the time of the marriage based on to the income shocks after the natural disasters: marry children at an earlier age if marriages favor their financial condition and marry children at a later age if marriage exacerbate the household financial position after natural disasters (Khanna & Kochhar, 2021). However, although the current theoretical model assume that natural disasters induced income shock has more pronounced effect among households with girls than household with boys as boys face a more inelastic income, it is not necessary true in any types of natural disasters (Corno, Hildebrandt, and Voena (2020). To better understand how earthquakes affect child marriages in Indonesia, we examine whether the direction of marriage payments determines how child marriage responds to earthquakes using variation in the practice of bride price across ethnic groups in Indonesia. The estimated impact of earthquakes on child marriage is statistically different for ethnic groups that practice bride price and ethnic groups that do not in the rural area. In the urban area, girls who are from bride price tradition group faces 2.9 percentage points increase in annual hazard entering child marriage compared those who do not practice bride price tradition. On the contrary, rural girls have 6.73 percentage points lower annual hazard entering child marriage, counting for about 185% reduction from the group average, if they are from bride price tradition group. Those results suggest that marriage payment influences child marriage in rural and urban Indonesia differently and the change in the urban area follows the theoretical prediction in Corno, Hildebrandt, and Voena (2020), but not in the urban area. Lack of more detailed dowry data and marriage payment data and annual household income data does not allow us to use empirical tools to verify the reasons why we observed such effect.

This paper contributes to two strands of literature in economics. First, it adds to studies that examine the impact of natural disasters in developing countries. Previous research has shown that natural disasters reduce human capital for individuals who are directly affected and their offspring (Caruso, 2017; Caruso and Miller, 2015), increase fertility and reduce child spacing (Nandi, Mazumdar and Behrman 2017), and age at marriage (Das and Dasgupta 2020; Khanna & Kochhar, 2021). Additionally, experimental evidence has demonstrated that natural disasters change trust, risk aversion, and patience (Cassar et al., 2017). In the long term, the literature has estimated positive effects of natural disasters on economic growth and welfare in Indonesia (Gignoux and Menendez, 2016; Heger and Neumayer, 2019). We contribute to this literature by focusing on the impact of natural disasters on child marriage. In a concurrent paper, Blanco (2023) studies the effect of earthquake on female marriage below the age of 18 in Indonesia using IFLS and USGS earthquake data. Our paper differs from hers in the follow ways. We employ detailed personal level migration history at specific subdistrict level, helping us precisely identify the location and time of each observation at the time of the earthquake. Second, we link the USGS ShakeMap data with our subdistricts level, which more accurately identifies the exposure to an earthquake at a more precise geographic unit. Next, we extract all 1,921 earthquakes that happened within Indonesia during the study period that could have impacted individuals at the subdistrict level, whereas she profiled only 21 earthquakes from 1994 to 2014. Lastly, we did not follow a difference-in-difference study design which does not require us to make parallel trends assumption for our analysis.

Second, this paper contributes to the literature on the causal determinants of child marriage. A strand of this literature has examined the impact of interventions using randomized controlled trials. Child marriage declined in response to unconditional cash transfers (Baird et al. 2011), financial incentives conditional on marriage (Buchmann et al., 2018), education subsidies (Duflo, Dupas and Kremer, 2015), and empowerment programs (Bandiera et al., 2020). Another strand of literature has studied the impact of government policies that target child marriage directly, for example, by increasing the legal minimum age for marriage or indirectly by increasing education. McGavock (2021) shows that child marriage declined after the legal minimum age for marriage was increased from 15 to 18 in Ethiopia. Breierova and Duflo (2004) show that a school construction program increased the age of marriage in Indonesia, while Keats (2018) and Chicoine (2020) show that exposure to free primary education led women to delay marriage. Besides the papers above on the impact of droughts, research has examined the effect of conflict on age at marriage, without looking at child marriage specifically (Jayaraman, Gebreselassie, and Chandrasekhar 2009; Shemyakina 2013). We contribute to this literature by shedding light on how natural disasters such as earthquakes affect child marriage.

The rest of the paper is as follows. Section 2 explores the conceptual framework of how earthquake might have affected the child marriage. Section 3 demonstrates the datasets and data cleaning process. Section 4 describes the estimation strategy and Sections 5 discusses the results, followed by a conclusion section.

Conceptual Framework

Natural disasters may affect child marriage through different channels. This section of the paper explores possible mechanisms based on the current literature. Marriage payment and consumption smoothing have been studied extensively in the literature (Corno, Hildebrandt, &

Voena, 2020). In the regions where marriage payment is prevalent, households utilize marriage payment to cope with negative income shock. In places where bride price is practiced, or money is transferred from the groom family to the bride family, natural disasters would increase child marriage among girls and decrease in boys as the household would marry daughters at an earlier age to cope with the income loss during natural disasters. On the other hand, in places where the dowry tradition is practiced, or money is transferred from the bride family to the groom family, we would observe the opposite effect (Das & Dasgupta, 2020; Khanna & Kochhar, 2021). Following the current theory, at the time of natural disasters, the consumption smoothing channel predicts that child marriage become prevalent among girls and less prevalent among boys among ethnic groups practicing bridle price in Indonesia, and no differential effect among ethnicities that do not follow marriage payment practice.

School and education could have affected child marriage after the natural disasters. The existing literature indicates that school produces a protective effect on child marriage as schooling delays sexual interaction with others (Keats, 2018), empowers children to make marriage decision by themselves (Jejeebhoy et al., 2012), and do not produce additional financial burden to the household (Khanna & Kochhar, 2021; Kumala Dewi & Dartanto, 2018). In the case of natural disasters, once schools are damaged or students are forced out of school after natural disasters, children might lose all those protective factors and get married at a younger age. Moreover, natural disasters induced education loss might lead to decreasing independence and social awareness of affected girls, gapping them from the protective effect of education on child marriage and forcing them to follow social norms to marry at young ages (Rumble et al., 2018). As such, earthquake-induced school dropout would impose a negative effect on girl and generate more child marriage.

The destructive effect of natural disasters may change the marriage market, imposing more negative impacts on girls and forcing the parents to marry them at earlier ages. Specifically, natural disasters might induce large displacement and social turbulences, all of which would increase insecurity and vulnerability among affected households (El Arab & Sagbakken, 2019; Blanco, 2023). Households with young age girls might face the largest impact as their daughters might be exposed to extreme events, such as rape or violence, and become unmarriable afterward. To protect the honor of the family, and their daughter's safety and seek protection, the best strategy is to marry their daughter after the natural disasters (United Nations Office for Economic and Social Commission for Western Asia, 2020; El Arab & Sagbakken, 2019; Buchmann et al., 2021; Kumala Dewi & Dartanto, 2018). In this case, earthquakes would increase child marriages, especially the marriage of young girls after earthquakes.

As natural disasters have disproportional risks to women, a sex ratio change with more man and fewer women are often observed (Neumayer & Plumper, 2007). Such a sex ratio change might lead to a marriage market change. With more men and fewer women, girls might have more power to choose better patterns for marriage (Khanna & Kochhar, 2021). However, the actual effect, whether they will delay the marriage for a better choice or marry at younger ages due to larger market demand, reminds unknown. Therefore, a change in the sex ratio cannot predict the effect of earthquakes on child marriage.

Lastly, natural disasters could also change marriage market search costs and delay marriage after natural disasters. As natural disasters disrupt daily life and economic production, individuals who start the marriage partner's searching process might delay their search, which eventually delays marriage and decrease child marriage afterward (Vogl, 2013). In addition, natural disasters induced temporary employment loss or financial instability might mask the true marriage market conditions, prolonging the partner-seeking process and decreasing child marriage (Khanna & Kochhar, 2021). Given those conditions, earthquakes could delay the timing of marriages and decrease child marriage. This paper examines the first two mechanisms.

Data

To estimate the impact of earthquakes on child marriage, we combine the Indonesian Family Life Survey (IFLS) with local measures of earthquake intensity from USGS. We use IFLS to obtain personal information on age at first marriage, ethnicity, religion, education, migration history, and household characteristics from all surveyed provinces in Indonesia. IFLS is an ongoing longitudinal survey tracking individuals over time. It represents 83% of the Indonesian population and contains over 30,000 individuals living in 13 of the 27 provinces in the country (Strauss, Witoelar, & Sikoki 2016). IFLS tracks complete marriage for all married individuals, including their marriage timing and marriage age, and the longitudinal structure of the survey and the detailed information on migration allows us to construct individuals' migration history, including the residential location at birth, at the age of 12, and any migration history afterward. As the IFLS only includes individuals older than 15 and the first wave was conducted in 1993 and the last wave in 2014/2015, we restrict the sample to the birth years from 1978 to 1996, in which the individuals would be at least 15 years old by the time of the first survey to fill out the marriage information survey and at least 18 years old in 2014 to track all possible age exposure to earthquakes[12]. After merging personal information and residential location, we obtain a longitudinal data set where the observation unit is a person-month of age,

[12] Indonesian Family Life Survey only starts to collect marriage information at age of 15 and any age before 15 would be considered as children where the marriage information is not asked. Anyone younger than 18 at the time of the last survey is dropped from the sample since we do not know their complete marriage information after the year of 2014.

from the first month of the 12th year of age to the last month of the 17th year of age. As the migration history before the age of 12 is not clarified, we restrict our sample starting age from the age of 12.[13] The residential locations are time-varying, which can take a different value each month.

We create a time-invariant variable that measures age at the first marriage based on self-reported marriage information. The first marriage time is reported by each individual. Some of them report accurate information include the first marriage year, month, and age. But some of them only report, first marriage year, or first marriage age, and some of them did not report their first marriage information even if they are married. Since not all individuals have complete information on the month of the marriage, we aggregate the personal information to yearly level and report the age of marriage in years at each age. More detailed information on the personal and migration data cleaning process can be found in Appendix A. In total, we obtained 19,180 individuals who were either not married yet at the time of the survey or have an age at first marriage, of which about 5% of married individuals have no age at first marriage and are, therefore, dropped from our sample. After merging migration history with marriage data, we obtained 17,397 individuals with complete marriage information and migration history, accounting for about 86% of the 19,180 individual sample with first marriage age.

In addition, we also construct before-marriage household characteristics to account for the influence of the family conditions on marriage decisions, such as parental education level, number of unmarried children within the household, the birth order of the children, and the family's primary religion and ethnicity. As socioeconomic conditions could determine how

[13] Marrying before age of 12 is also a rare event. Only 63 married individuals reported they married before age of 12 in our sample.

earthquakes affect marriage decisions, we construct a Principal Component Analysis (PCA) following Vyas and Kumaranayake (2006) by including three vectors of household variables: household infrastructure, durable household assets, and household living conditions. Lastly, we create a quintile of the first component of the PCA to indicate the socioeconomic conditions of the households where the first quantile indicates the poorest and the fifth quantile indicates the richest household at the time of each survey (McKenzie, 2005). One critical household characteristic variable used in this paper is the residential locality of the individual. The head of the household reports this time-variant variable at the time of the survey. We followed the previous study and assigned the first available household residential locality information from ages 12 to 17 to individuals throughout the study period at each age year (Gignoux and Menéndez 2016). As we only observe the household characteristics at the time of each survey, we use the panel information from the last survey to fill out the information until the following survey is conducted. Since earthquakes could have impacted the household conditions, we limit the household information to before-marriage information. Their household characteristics information will be used if someone was already married at the time of the survey. More details of the household characteristics can be found in Appendix A.

For earthquake information, we obtained ShakeMap data from the United States Geology Survey (USGS) from 1980 to 2015 within the vicinity of [-12, 12] Latitude and [80, 150] Longitude.[14] ShakeMap combines ground motion models and recorded data to create a ground shaking map showing the intensity of the earthquakes by Modified Mercalli Intensity (MMI). ShakeMaps are made in real-time by automated systems, restored in ComCat (short for

[14] The selected time range is arbitrary, but the year 2015 is the latest IFLS5 years. Therefore, we cannot study the effect of the earthquake after 2015, and the select map area is larger than the Indonesia boundary to ensure all earthquakes are selected. If an earthquake is not located within the Indonesia boundary, no spatial link between subdistrict and earthquake would be found.

Comprehensive Catalog), and linked to other products associated with the source earthquake (PAGER, DYFI, Finite Fault, etc.). However, ShakeMaps only record earthquakes in any affected area, with MMI higher than 5.5 for global events and 3.5 for US domestic events.[15] Compared to the Gignoux and Menéndez (2016) using earthquake information to estimate the peak group acceleration (PGA) and MMI, this new data provides more accurate information on earthquake intensity.[16] MMI tracks the group shaking intensity based on individual reports of damages and perceived shaking for each earthquake, ranging from the Roman number I to XII, and the intensity increases with numbers, providing a measurement of the earthquake effect at small geographic units of each polygon (USGS, 2022; Wald et al. 1999).

To link the earthquake's impact with each individual, we count on the migration history of the individual from the age of 12 each month. Individuals report their place of birth, place of residence when 12 years old, and any migration afterward at four different administrative levels: province (provinsi), district (kabupaten), subdistrict (kecamatan), and village (kelurahan). However, to protect the privacy of the surveyed individuals, only the first three administrative levels are reported, and those administrative units are recorded by Badan Pusat Statistik using BPS code in the format of XX-YY-ZZZ. XX presents the province code, YY records the district code, and ZZZ indicates the subdistrict code. As the boundary of the subdistricts could change over time due to merging, splitting, and changing of subdistricts boundaries, BPS codes change

[15] We appreciate the support and help from USGS staff on helping us understand how ShakeMap products are recorded. Hearne, Mike <mhearne@usgs.gov> provided detailed ShakeMap explanation to us, which we are not able to be found online.

[16] Gignoux and Menéndez (2016) uses Centennial Earthquake Catalog data to estimate the Peak Ground Acceleration (PGA) and Modify Mercalli Intensity (MMI), which has been replaced by new dataset provided by USGS at https://earthquake.usgs.gov/earthquakes/search/. We used ComCat to download all selected earthquakes using Python by inputting [-12, 12] Latitude and [80, 150] Longitude from 0:00:00 am January 1st, 1980, to 11:59:59 pm December 31st, 2014. ShakeMap provides earthquake MMI shapefiles, mapping the earthquake intensity for each earthquake recoded by polygons. Therefore, no estimation on MMI distribution is needed.

over time. Therefore, we combined all changed boundary subdistricts into redefined subdistricts using the Crosswalk provided by IFLS. This process ensures that the unit of subdistricts is consistent across all survey waves. IFLS does not provide the spatial distribution of subdistricts, which maps the subdistricts and their geographic locations on map for geoprocessing. Thus, we connect the IFLS subdistricts through another dataset, GADM, to obtain their spatial distribution by using the BPS code.[17] If the subdistrict had merged, split, or changed throughout survey time, we would merge those linked subdistricts as one subdistrict. A cut version of before and after merge maps are illustrated in Figure 2, and an example of the dissolving subdistricts process is shown in Appendix A.

Lastly, we spatially connected subdistricts and earthquake MMI distribution, taking the arithmetic average of all connected MMI polygons within that subdistrict. By merging USGS and GADM, we have the average MMI of each subdistrict at the time of each earthquake. Figure 3 provides an example of this process. As our smallest time unit is a month and the smallest geographic unit is a subdistrict, multiple earthquakes could happen in the same subdistrict and same month. Therefore, if such multiple events happened within a month, we only keep the one with the largest MMI within that subdistrict month. Finally, we merge the earthquake information and personal data by residential location (BPS code) and time (year and month). However, since we cannot accurately identify the marriage time in months, we collapse the data from person month to person year. Then, we define the exposed age as that individual who experienced an earthquake at that age year during any month from age 12 to 17.[18] If an

[17] GADM was published in 2009, and it provides spatial data for all countries and their sub-divisions.

[18] This process could create some measurement error. For example, if someone experienced the earthquake at 13 years and 11 months old. We treat the exposed age as 13 years old, but the earthquake effect should reside in age of 14. As we do not have accurate information on marriage month, we cannot accurately identify the marriage timing in months.

individual experiences multiple earthquakes within that year, we keep the highest MMI value of that age year. More details of this process can be found in Appendix A. Thus, we have the individual information from 12 to 17, recording their data and an indicator if this individual experienced any earthquakes at a certain age. Table 1 summarizes the data for this study.

Estimation

Our identification strategy relies on the quasi-random nature of earthquakes in Indonesia (Gignoux and Menendez 2016). To examine the impact of earthquakes on the timing of marriage, we use a simple discrete approximation duration model, following Corno, Hildebrandt, and Voena (2020). The duration of interest is the survival time between t_0, the age of an individual who first faces the risk of getting married, and t_m, the time of first marriage. We choose t_0 to be the age of 12 as IFLS records the residential location from the age of 12 and onward, and t_m is the age at first marriage. We aggregate our data to the year level and create a person-age year panel data format, observing their marriage decision, the experience of earthquakes, and other personal information since the age of 12 at each age year. Moreover, since we study how earthquakes affect child marriage, marriage decision after the age of 17 is out of the scope of this study, so our observation period is from the age of 12 to 17. Each observation is a person-year from 12 until they are married or more than 17 years old. We merge this individual information with earthquake information. As the earthquake is recorded monthly, we define an individual who experienced an earthquake if at least one earthquake happened in that age year. If multiple earthquakes occurred at the same age year, we only record the earthquake with the largest MMI, but it is very rare to have two earthquakes happening within the same year. The following specification demonstrates the probability model used for our study,

$$M_{isyt} = \beta E_{syt} + X_{syt} + \alpha_t + \gamma_s + \mu_y + \epsilon_{isyt}$$

M_{isyt} is a binary variable, equal to 1 in the year the individual is married and zero otherwise for each individual i, subdistrict s, year y, and age t. As we proximate the duration model and study the hazard of the earthquake on child marriage, the age years after the first marriage year are right-censored. The variable E_{syt} is a time-variant binary indicator for exposure to an earthquake with MMI VI (6) or higher in that age year. We choose VI as the cutoff as intensities at this level generate non-negligible damage to vulnerable and resistible buildings (Wald et al. 1999). A robustness of choosing MMI V and VII are conducted and presented in Appendix B.1. β is the coefficient of interest, and it measures the hazard of earthquakes on marriage before the age of 17 at any given age if they are not married yet. X_{syt} is a vector of control variables, measuring individual and household characteristics. α_t is the age fixed effect, recording the different probability of marriage at each given age during the study age period. In addition, γ_s is the subdistrict fixed effect, controlling the unobservable subdistrict characteristics that do not change over time but could have affected the individual probability of being married. Finally, μ_y is the cohort fixed effect, tracking the time-variant factors that could have affected the individual marriage time. By incorporating the cohort fixed effect and subdistrict fixed effect, earthquake's impact on the child marriage hazard is identified within location and within year of birth in earthquakes and marriage time.

To verify the identification strategy, we explore the earthquake's leading effect before the earthquake happened. Had the earthquake been unpredictable, we would not have observed the effect of earthquakes on child marriage and other associated outcomes before the earthquake. The leading effect results indicate that the effect of earthquake decreased to 0.7 percentage points level, counting for 50% reduction in group mean and significant at 0.1 level, in the urban area (see Table B.2 in Appendix B). The effect diminished to insignificant level in the rural area despite a 57% increase in the annual hazard rate in the rural area. We argue that since the age, birth year, and marriage time are not accurately reported, measurement error in timing might explain those effects before the earthquake happened. More details will be discussed in Appendix B.

Results

Main Results

Table 2 reports our main results on the effect of earthquakes on the timing of marriage in the exposed age year for both men and women from the age of 12 to 17. Panel A reports estimates for girls and Panel B shows results for boys. Columns 1 to 3 identify the effect of the earthquake using three different specifications. The first columns only include age fixed effect, birthyear fixed effect, and subdistricts fixed effect. The second columns include personal control variables, including ethnicity and religion dummies. The third columns include household control variables, including household wealth quantile, order of birth within the family, number of children within the family, and parental education dummies for less than primary, less than compulsory, or finished compulsory or higher with less than primary is dropped from regression as reference group. By adding household variables, the sample size was reduced significantly since having household characteristics requires the individual has panel information from the last survey. To account for potential bias and sample change, we run regression without personal and household variables but use the same sample with household and personal variables. We observe similar results in this specification, but this result is not reported in the table. The estimated coefficients are relatively small and imprecisely estimated for all individuals, suggesting that there is no overall effect of earthquakes on child marriage. However, this overall effects masks heterogeneity by residence. A subgroup analysis of urban and rural residents shows that urban girls who experienced an earthquake had MMI higher than VI between ages 12 and 17 are about 1.24 percentage points (pp) less likely to get married in the same year. This result is consistent across different specifications (columns 3-6, p<0. 01). The average annual marriage hazard for this age group is equal to 0.0128, and hence the effect corresponds approximately to a 100% decrease (83% decrease if counting for household control variables). In contrast, rural girls who experienced an earthquake with MMI higher than VI are about 5 to 7 pp more likely to get married in the same year (columns 7-9, p<0. 05), showing about 200% increase in the annual marriage hazard. The Chow Test indicates a significant difference in the annual marriage hazard for girls between rural and urban groups (F-tests are 17,44, 16.86, and 3,24 p-value less than 0.01, 0.01, and 0.05 respectively for three specifications). As for boys, although the estimated effect is relatively large, for example, the annual marriage hazard increases by about 679% for boys living in the urban area at the time of the earthquake, the coefficient is not precisely estimated at the conventional level. The Chow test also shows a significant difference in coefficient between boys who lived in rural and urban areas. Notably, in the effect of earthquakes, urban and rural annual marriage hazard changes in the opposite direction among boys and girls. This difference might indicate the change of the marriage market equilibrium, which requires future investigation.

We further explore the heterogeneity of this effect by the girl's age by interacting earthquake with each age dummy. Table 3 shows the earthquake effects in different ages between girls and boys when controlling only for the personal control variables, and Figure 2 plots the estimated effect of earthquakes on the annual hazard into marriage by age group.[19] Overall, for girls who are at the age of 12, 13, and 14, they face a 1.62 pp, 1.63 pp, and 0.89 pp increases in the annual marriage hazard in the event of earthquakes ($p<0.01$). Although the annual marriage hazard still increases at the ages 15 and 16, those effects are not imprecisely estimated ($p>0.1$). Once the girl is 17 years old, the event of an earthquake decreases their annual marriage hazard by 6.3 pp, which is a 225% reduction relative to the sample average. Urban and rural samples have similar results as all individuals, with the exception that urban girls start to have a reduction in marriage hazard starting at the age of 15 and no significant effect at the age of 12, and rural girls experience insignificant positive hazard at the age of 14. Similar to the main results, the earthquakes did not generate a statistical effect for boys ages 12 to 16 but resulted in a large reduction in marriage hazards at age 17 regardless of whether they lived in the rural or urban areas.

As we cannot precisely identify the time of marriage, measurement error could result in a one-year marriage age difference in the reported marriage age.[20] Moreover, the effect of the earthquake might linger after the earthquake due to it large destructive power, resulting in a lagged effect in the following years. To account for those two considerations, we estimate the first lag of the earthquake effect on marriage. Since one individual could experience more than one earthquake during the study time period, studying lagged effect could have inaccurate results if another earthquake happened one year after the first earthquake or had multiple earthquakes after. However, consecutive earthquake events are rare in our dataset. Therefore, we dropped

[19] Only regression results controlling for fixed effects and personal control variables are reported in the table. But the result is consistent across different specificity when control for only fixed effects and household variables.
[20] For example, when someone reported they married at age of 14, but in fact it was age of 13 and 11 months. If the earthquake happened at age of 13, we would not capture the true earthquake effect at age of 13. Lagging earthquake to age of 14 would help to capture the effect of earthquake from age of 13 to age of 14.

those who had more than one earthquake during the study period from our sample, restricting one person from one earthquake condition for this analysis. The estimated result is demonstrated in Table 4 Panel A. In addition, as we do not observe the residential location at the age of 11, we do not know if individuals experienced earthquakes at 11. Thus, we cannot know what happened at the age of 12 when utilizing the first lag model. To account for potential bias, we first dropped the age of 12 and studied the first lag model using ages 13 to 17, and the results are shown in Table 4 Panel A. To preserve the sample and ensure the earthquake effect at the age of 12 is preserved, we assume that no earthquakes happened or that the lagged effect from the age of 11 to the age of 12 is minimal. Under this assumption, we preserve the sample and rerun the same regression. The result in shown in Table 4, Panel B. The estimated effect of experiencing an earthquake one year before does not have a robust pattern, and we cannot conclude that the earthquakes had significantly lagged effects on marriage one year after.

Marriage Payment

One possible reason earthquakes could have affected the probability of entering child marriage is the marriage payment system practiced in many developing countries. Marriage payments take two forms: bride price and dowry system. Under the bride price system, the groom's family sends money or monetary assets in exchange for marriage (Corno, Hildebrandt, and Voena, 2020). After extreme events such as natural disasters, when the bride's families face assets loss or monetary assets, they would be more prone to marry their daughters at an earlier age in exchange for monetary assets, leading to more child marriage. On the other hand, under the dowry system, the bride's family prepares for the dowry and uses it to ensure their daughter marries someone with matched social and economic background (Das and Dasgupta, 2020;

Khanna & Kochhar, 2021). Similarly, the bride's families would delay their daughters as they need more time to prepare for dowries, resulting in less child marriage.

There are many ethnicities in Indonesia, and different ethnicities may practice different marriage payments system. We follow Ashraf et al. (2020) to separate all IFLS ethnicities into two groups: bride price and non-bride price ethnicities.[21] Following the theoretical framework from Corno, Hildebrandt, and Voena (2020), If the bride price had played a role after the earthquake, the destructive effect of the earthquake would have increased the child marriage probability in the bride price group as the bride families need monetary assets after earthquakes to cope with the financial loss. Based on the results from Table 5, we do not observe a significant effect of the bride price indicator in the regression. The interaction between earthquake and bride price indicators shows that girls living in the urban area have about 2.9 pp higher annual marriage hazard rate, and girls living in the rural area have about 7 to 9 pp lower marriage hazard rate if there are from bride price tradition group. The interaction coefficient changes in the opposite direction with the main earthquake effect. There is no significant effect observed among boys.

This results partially follows the theoretical indications from the previous research on the effect of marriage payment. Corno, Hildebrandt, and Voena (2020) predicts that for groups that practice bride price, earthquake induced income shock would increase the annual hazard ratio for girls to marry at earlier age. We observe a different effect of earthquakes on marriage rates among ethnical groups that practices bride price from our results where only urban not rural groups followed the theoretical prediction. One explanation is that Corno, Hildebrandt, and

[21]Bride price ethnicities are Batak; Bugis, Madura, Sasak, Banjar, Makassar, Nias, Palembang, Toraja, Betawi, Melayu, Ambon, Manado. The ethnicities Chinese, Sumbawa, Dayak, Other South Sumatra, Banten, Cirebon, Other are not coded in the Ashraf et al. (2020). We code them as not practicing bride price.

Voena (2020) assume that income elasticity is higher in the grooms and their families and lower in brides and their families in the case of extreme events. However, this assumption might not be held in our case, in which boys face similar financial loss as girls and cannot afford marriage payments afterward. The absence of the annual or more frequent household income or individual work data leaves this statement untestable.

Education

Education has been proved to induce a protective effect on child marriage (Paul 2019; Raj et al. 2019). This section explores how education or being in school at the time of the earthquake could have modified the results. Considering earthquakes could also affect the probability of an individual staying in school (Rush 2018), we add education status from the previous year to ensure education status enters the regression without creating an endogeneity issue between education and marriage. Table 6 shows the result using the same exact specification as before but adding education variables. First, we observe a consistent effect for girls who were in school the year before the earthquake, the annual marriage hazard rate decreased by 0.7 pp in the total sample and in the rural subsample, but not for the urban sample. This result indicates that education may only provide a protective effect for girls in rural areas but in urban areas. The result is similar for boys, with no significant protective effect observed for those living in the urban area. The interaction between the earthquake and in-school status from last year shows that being in school the year before the earthquake increases the marriage hazard by 3 pp in the rural area but decreases the marriage hazard by 3 pp in the urban area. However, both interaction terms are not significant at conventional level, indicating the difference between two groups are likely to be random. Interestingly, the interaction terms change in the opposite direction in the urban and rural areas, following the same indication of bride price interaction term where the effect of earthquakes had fundamentally different effects in the urban and rural areas.

Table 7 presents the results of the earthquake's effect on school dropouts as the outcome. Even though the earthquake negatively affects education continuation for girls, those effects are imprecisely estimated. A weak protective effect is observed for boys living in rural areas, but statistically insignificant.

First Pregnancy

Pregnancy is directly related to marriage, and early pregnancy could negatively impact the mother's health and their offspring (Xu et al. 2003; Nove et al., 2014). Following the previous study design, we explore the effect of the earthquake on pregnancy from the age of 12 to 17. As IFLS only reports pregnancy history among married women, we restrict our samples to married women and drop all unmarried women in this section. In addition, as IFLS only records the time of the delivery or miscarriage, not the time of conceiving babies, we add a lagged earthquake term from the previous year to count for this time difference. Table 8 reports the results. Overall, earthquakes decrease the annual hazard rate of pregnancy by 3 pp in the current year and 2 to 3 pp in the year following the earthquake among urban married women when we control for fixed effects and personal variables, but not household variables. This change is mostly likely to be explained by the sample size drop, as not every woman could have before-marriage household characteristics information in the sample. We observe that earthquakes increase the annual hazard rate of pregnancy by about 2 pp in the current year and 1 to 3 pp in the following year for married girls who live in the rural area, but those effects are imprecisely estimated.

Conclusion

This paper examines the impact of natural disasters on child marriage in developing countries. First, we study the effect of earthquakes in Indonesia, a country with high seismological activity and a high prevalence of child marriage. Our results reveal that women who experience an earthquake between ages 12 and 17 are 1 pp less likely to get married in the same year if they lived in urban areas and 5 to 7 pp more likely to get married in the same year if they lived in urban areas. The magnitude of the effect is very large, counting for about 100% reduction in sample average in urban area and 137% increase in the rural area. Boys face opposite effects in two areas, giving about 0.7 pp increase in urban and 0.5 pp decrease in the rural area. However, the effect for boys is imprecisely estimated. This result implies that earthquakes have a significant heterogeneous effect on rural and urban girls. In addition, extremely young age girls suffer the most from earthquakes and face the highest marriage hazard rate in the event of an earthquake, but we do not observe a significant and robust effect for boys. Second, to explore the mechanism of the earthquake effect, we exploited sub-national variation in the direction of marriage payments to test whether the impact of earthquakes on child marriage varies across groups that practice bride price. We found that the bride price tradition suggests that consumption smoothing affects the timing of the marriage in the case of earthquakes differently in the urban and rural areas, but we cannot testify mechanism behind the effects due to the data limitation. Third, we did not find being in school the year before the earthquake generate any statistically significant differential effect on marriage hazard ratio compared to those who were not in school the year before the earthquake. Lastly, being exposed to earthquakes delay the first pregnancy among married women, and this effect is more pronounced among women living in the urban area.

Tables & Figures

Table 16: Summary Statistics

	Boys			Girls		
	N	Mean	St. dev.	N	Mean	St. dev.
Panel A: Individual data						
Age of first marriage conditional on marriage	4,503	23.68	3.684	6,758	20.53	3.785
% Married between ages 12 and 17	8,062	0.021	0.142	8,963	0.168	0.374
Ethnicity is Javanese	7,924	0.426	0.495	8,809	0.427	0.494
Ethnicity is Sundanese	7,924	0.120	0.326	8,809	0.117	0.322
Ethnicity is Balinese	7,924	0.043	0.203	8,809	.0437	0.204
Religion is Muslim	8,020	0.893	0.309	8,895	0.897	0 .303
Religion is Protestant	8,020	0.046	0.209	8,895	0.041	0.199
Religion is Hindu	8,020	0.043	0.203	8,895	0.043	0.205
Number of earthquakes with MMI greater than VI	8,062	0.027	0.163	8,963	0.022	0.149
Ever experienced an earthquake with MMI greater than VI	8,062	0.026	0.160	8,963	0.021	0.143
Panel B: Survival data						
Age	48,372	14.5	1.708	53,778	14.5	1.708
% Age-year with MMI greater than VI	48,372	0.004	0.067	53,778	0.003	0.060
Lives in an urban area	32,903	0.540	0.498	32,516	0.505	0.500
Wealth Quantile (1=poorest and 5=richest)	32,900	2.915	1.398	32,508	2.848	1.381
Order birth within the household (1=first born)	27,960	1.672	0.965	28,439	1.804	1.019
Number of unmarried siblings	27,960	0.673	0.965	28,439	0.805	1.019
Number of children living within the household	27,960	2.866	1.447	28,439	3.026	1.496
Father had less than primary education	22,892	0.377	0.485	24,503	0.387	0.487
Father had less than compulsory education	22,892	0.292	0.455	24,503	0.294	0.456
Father had compulsory or higher education	22,892	0.331	0.471	24,503	0.318	0.466
Mother had less than primary education	26,141	0.462	0.499	27,298	0.471	0.499
Mother had less than compulsory education	26,141	0.301	0.458	27,298	0.297	0.457
Mother had compulsory or higher education	26,141	0.238	0.426	27,298	0.232	0.422

Note: Panel A summarizes information at each individual level, and Panel B summarizes the data at the person-year level. The list of the ethnicity only lists the top three most prevalent ethnicities. Batak, Sasak, Minang, and Betawi are not listed in this table.

Table 17 : The Effect of Earthquakes on the Annual Hazard of Child Marriage

Panel A: Girls

	All Individuals			**Urban**			**Rural**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake	-0.0014	-0.0012	0.0190	-0.0124***	-0.0124***	-0.0115***	0.0502**	0.0507**	0.0738**
	(0.0077)	(0.0076)	(0.0146)	(0.0038)	(0.0034)	(0.0035)	(0.0240)	(0.0235)	(0.0362)
Observations	52,133	51,197	21,795	17,256	17,006	10,497	17,130	16,888	11,298
R-squared	0.030	0.030	0.029	0.019	0.020	0.026	0.038	0.039	0.039
N of subdistricts	1,520	1,515	631	602	598	395	515	508	378
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal Var	-	X	X	-	X	X	-	X	X
Household Var	-	-	X	-	-	X	-	-	X
Mean of Y	0.0288	0.0281	0.0252	0.0128	0.0124	0.0137	0.0364	0.0365	0.0359

Panel B: Boys

	All Individuals			**Urban**			**Rural**		
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Earthquake	0.0040	0.0037	0.0098	0.0076	0.0076	0.0141	-0.0049	-0.0051	-0.0034
	(0.0042)	(0.0041)	(0.0091)	(0.0063)	(0.0062)	(0.0120)	(0.0045)	(0.0045)	(0.0032)
Observations	48,212	47,235	20,819	18,782	18,473	10,411	16,492	16,168	10,408
R-squared	0.005	0.005	0.005	0.002	0.002	0.007	0.008	0.009	0.010
N of subdistricts	1,369	1,365	630	611	609	394	514	511	361
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal Var	-	X	X	-	X	X	-	X	X
Household Var	-	-	X	-	-	X	-	-	X
Mean of Y	0.0034	0.0033	0.0024	0.0011	0.0010	0.0009	0.0038	0.0038	0.0039

All regressions include age fixed effect, subdistrict fixed effect, and birth year fixed effect. Observations are in individual age structure, in which each observation counts for one individual from the age of 12 until the age of first marriage or age of 17. Personal controls include ethnicity and religion, and all are entered as dummies. Household controls are wealth quantile, order of birth within the family, number of children within the family, and parental education, entered by dummies of less than primary, less than compulsory, or finished compulsory or higher. By adding household variables, the sample size was reduced significantly. To count for potential bias and sample change, we run regression without personal and household variables using the same sample with household and personal variables. We observe similar results using the sample with personal and household variables. The Chow Test indicates a significant different annual marriage hazard between rural and urban girls, but no difference between boys.

*** p<0.01, ** p<0.05, * p<0.1

Table 18: The Effect of Earthquakes on the Annual Hazard of Child Marriage by Ages

	Girls			Boys		
	All Individuals	**Urban**	**Rural**	**All Individuals**	**Urban**	**Rural**
	(1)	(2)	(3)	(4)	(5)	(6)
Earthquake X Age of 12	0.0162***	-0.0100	0.0433***	0.00139	-0.000731	0.00294
	(0.0037)	(0.0098)	(0.0094)	(0.0013)	(0.0010)	(0.0022)
Earthquake X Age of 13	0.0163***	0.0055*	0.0270***	0.000992	0.0014***	-0.0001
	(0.0036)	(0.0028)	(0.00939)	(0.00123)	(0.0005)	(0.0040)
Earthquake X Age of 14	0.0089***	0.0077**	0.00907	0.000507	0.0003	-0.0135
	(0.0034)	(0.0035)	(0.00977)	(0.00117)	(0.0008)	(0.0186)
Earthquake X Age of 15	0.0074	-0.0245**	0.108	0.000544	0.000153	0.000294
	(0.0255)	(0.0111)	(0.119)	(0.000850)	(0.000783)	(0.00143)
Earthquake X Age of 16	0.0121	-0.0121***	0.115	0.0208	0.0392	-0.00167
	(0.0340)	(0.0042)	(0.0975)	(0.0203)	(0.0337)	(0.00143)
Earthquake X Age of 17	-0.0635***	-0.0441***	-0.0801***	-0.00846***	-0.00122*	-0.0119***
	(0.0042)	(0.0064)	(0.00786)	(0.00136)	(0.000707)	(0.00270)
Observations	51,197	17,006	16,888	47,235	18,473	16,168
R-squared	0.030	0.020	0.040	0.005	0.004	0.009
N of subdistricts	1,515	598	508	1,365	609	511
Fixed Effects	X	X	X	X	X	X
Personal Var	X	X	X	X	X	X
Household Var	-	-	-	-	-	-
Mean of Y	0.0281	0.0124	0.0365	0.0033	0.0010	0.0038

The regression specification includes all age fixed effect, subdistrict fixed effect, and birth year fixed effect and adds each age interacts with an earthquake. The coefficient indicates the differential effect of each age. Specifications with personal control variables are shown in the table, but the results are robust across all specifications when not including personal control variables and include household variables.

Table 19: The Lagged Effect of Earthquakes in the Following Year

Panel A: Drop age year at the age of 12

	All Individuals			Urban			Rural		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake t=0	-0.00610	-0.00625	0.0118	-0.0136***	-0.0139***	-0.0149***	0.0400	0.0396	0.0616*
	(0.00698)	(0.00688)	(0.0137)	(0.00522)	(0.00533)	(0.00510)	(0.0272)	(0.0258)	(0.0368)
Earthquake t=-1	-0.00209	-0.00280	-0.0158***	-0.00626	-0.00699	-0.0192***	0.0104	0.00905	-0.0219*
	(0.0108)	(0.0108)	(0.00466)	(0.0126)	(0.0122)	(0.00547)	(0.0256)	(0.0261)	(0.0120)
Constant	0.00330	0.000630	-0.000260	0.00283	-0.0113	-0.0246	0.00263	-0.00662	-0.00691
	(0.00443)	(0.00776)	(0.0131)	(0.00547)	(0.0105)	(0.0164)	(0.00866)	(0.0209)	(0.0259)
Observations	43,146	42,382	18,457	14,387	14,179	8,947	14,303	14,099	9,510
R-squared	0.026	0.026	0.026	0.017	0.018	0.025	0.031	0.032	0.035
N of subdistricts	1,505	1,500	626	599	595	391	527	520	375
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal	-	X	X	-	X	X	-	X	X
Household	-	-	X	-	-	X	-	-	X
Mean of Y	0.0342	0.0333	0.0294	0.0163	0.0158	0.0161	0.0469	0.0470	0.0420

Panel B: Assume no earthquake had happened or no earthquake effect at the age of 11

	All Individuals			Urban			Rural		
	(13)	(14)	(15)	(19)	(20)	(21)	(22)	(23)	(24)
Earthquake t=0	-0.00680	-0.00669	0.00936	-0.0139***	-0.0140***	-0.0130***	0.0350	0.0348	0.0538
	(0.00621)	(0.00616)	(0.0124)	(0.00414)	(0.00423)	(0.00380)	(0.0250)	(0.0242)	(0.0358)
Earthquake t=-1	-0.000631	-0.00103	-0.0142***	-0.00464	-0.00511	-0.0163***	0.00509	0.00434	-0.0172*
	(0.00846)	(0.00842)	(0.00368)	(0.00980)	(0.00949)	(0.00419)	(0.0210)	(0.0212)	(0.00899)
Constant	0.00313	0.00169	0.000834	0.00141	-0.00932	-0.0187	0.000730	0.00132	0.00551
	(0.00380)	(0.00658)	(0.0124)	(0.00458)	(0.00874)	(0.0153)	(0.00721)	(0.0193)	(0.0264)
Observations	52,104	51,168	21,778	17,307	17,053	10,497	17,338	17,091	11,281
R-squared	0.030	0.030	0.029	0.019	0.020	0.026	0.037	0.038	0.039
N of subdistricts	1,520	1,515	631	607	603	395	531	524	378
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal Var	-	X	X	-	X	X	-	X	X
Personal Var	-	-	X	-	-	X	-	-	X
Mean of Y	0.0288	0.0281	0.0252	0.0136	0.0132	0.0137	0.0397	0.0398	0.0359

Table 20: The Effect of Earthquakes on the Annual Hazard into Child Marriage Interact with Bride Price Practice

	All Individuals			**Urban**			**Rural**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake	-0.000479	-0.000464	0.0211	-0.0128***	-0.0127***	-0.0114***	0.0641**	0.0640**	0.0881**
	(0.00855)	(0.00854)	(0.0154)	(0.00386)	(0.00406)	(0.00351)	(0.0278)	(0.0273)	(0.0416)
Bride Price (BP)	-0.00126	-0.00444	0.0209	-0.00119	-0.00233	-0.0120	-0.00230	-0.00828	0.0177
	(0.00337)	(0.00733)	(0.0219)	(0.00436)	(0.00857)	(0.0113)	(0.0145)	(0.0290)	(0.0267)
Earthquake X BP	-0.00749	-0.00643	-0.0251	0.0288***	0.0290***		-0.0673**	-0.0650**	-0.0867**
	(0.00938)	(0.00947)	(0.0197)	(0.00589)	(0.00666)		(0.0293)	(0.0289)	(0.0439)
Observations	49,792	49,688	21,157	16,719	16,683	10,314	16,307	16,239	10,843
R-squared	0.030	0.030	0.028	0.020	0.020	0.026	0.038	0.039	0.039
N of subdistricts	1,504	1,503	623	586	586	392	497	496	368
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal Var	-	X	X	-	X	X	-	X	X
Household Var	-	-	X	-	-	X	-	-	X
Mean of Y	0.0279	0.0278	0.0248	0.0125	0.0125	0.0137	0.0362	0.0362	0.0354

All regressions follow the same specification as in table 2 except for adding two terms, Earthquake X BP and BP. We follow Ashraf et al. (2020) and categorize bride price (BP) as individuals identified as Batak, Bugis; Madura; Sasak, Nias, Toraja, and Ambon. For those not on the list, such as Chinese, Sumbawa, Dayak, Other South Sumatra, Banten, Cirebon, we categorized them as the non-bride price group. Specification 6 does not have an interaction between Earthquake and BP since there is not enough observation experience of earthquakes within the BP group in the urban setting with all household variables. Regression using the same sample with household and personal variables indicates similar results but did not report in the table. The table does not report regression results of boys, as no statistically significant effects are observed.
*** p<0.01, ** p<0.05, * p<0.1

.

Table 21: The Effect of Education on the Child Marriage Annual Hazard from Earthquakes

	All Girls			Urban Girls			Rural Girls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake	-0.00105	-0.00658		-0.0127***	-0.0112**		0.0494**	0.0307	
	(0.00784)	(0.0112)		(0.00418)	(0.00503)		(0.0233)	(0.0356)	
Earthquake X In School		0.0113	0.00469		-0.00334	-0.0145**		0.0335	0.0642*
		(0.0168)	(0.0116)		(0.00677)	(0.00572)		(0.0530)	(0.0344)
Earthquake X Out of School			-0.00761			-0.0184**			0.0377
			(0.0163)			(0.00768)			(0.0397)
Attending school in the last year	-0.0071***	-0.0072***	-0.0072***	-0.0030	-0.0030	-0.0030	-0.0071**	-0.0072**	-0.0072**
	(0.00149)	(0.00149)	(0.00149)	(0.00205)	(0.00207)	(0.00207)	(0.00303)	(0.00304)	(0.00303)
Observations	50,481	50,481	50,481	16,964	16,964	16,964	16,884	16,884	16,884
R-squared	0.031	0.031	0.031	0.019	0.019	0.019	0.038	0.038	0.038
N of subdistricts	1,499	1,499	1,499	596	596	596	522	522	522
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal Var	X	X	X	X	X	X	X	X	X
Household Var	-	-	-	-	-	-	-	-	-
Mean of Y	0.0279	0.0279	0.0279	0.0128	0.0128	0.0128	0.0397	0.0397	0.0397

All regressions control for fixed effect and personal variables and standard error is clustered at the subdistrict level. Colum (1), (4), and (7) uses the main specification and add one indicator if the individual was in school or not last year. Colum (2), (5), and (8) uses the main specification and add one education indicator as before and one interaction term between the earthquake and education indicator. Colum (3), (6), and (9) use the main specification and education indicator and out of school indicator and in school indicator in the last year to interact with an earthquake.
*** p<0.01, ** p<0.05, * p<0.1

Table 22: The Effect of Earthquakes on the Annual Hazard in School Dropouts

Panel A: Girls

	All Individuals			Urban			Rural		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake	-0.0202	-0.0221	-0.0645	-0.0520	-0.0525	-0.0585	-0.0377	-0.0377	-0.0899
	(0.0454)	(0.0454)	(0.0473)	(0.0584)	(0.0574)	(0.0685)	(0.0813)	(0.0852)	(0.0669)
Observations	31,999	31,533	13,700	11,318	11,182	6,983	9,901	9,752	6,717
R-squared	0.075	0.076	0.131	0.122	0.125	0.146	0.130	0.132	0.134
Number of BPS	1,428	1,423	556	524	523	358	409	404	308
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal Var	-	X	X	-	X	X	-	X	X
Household Var	-	-	X	-	-	X	-	-	X
Mean of Y	0.209	0.208	0.221	0.181	0.181	0.199	0.229	0.229	0.243

Panel B: Boys

	All Individuals			Urban			Rural		
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Earthquake	0.00239	0.000591	0.0360	-0.0174	-0.0185	0.00752	0.0204	0.0220	0.0463
	(0.0337)	(0.0336)	(0.0529)	(0.0411)	(0.0412)	(0.0549)	(0.0681)	(0.0683)	(0.118)
Observations	29,922	29,388	13,135	12,399	12,208	6,877	9,787	9,622	6,258
R-squared	0.091	0.091	0.135	0.135	0.136	0.139	0.149	0.150	0.150
Number of BPS	1,244	1,240	569	553	551	369	407	403	302
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal Var	-	X	X	-	X	X	-	X	X
Household Var	-	-	X	-	-	X	-	-	X
Mean of Y	0.200	0.200	0.210	0.186	0.186	0.200	0.208	0.207	0.222

All regressions control for fixed effect and personal variables and standard error is clustered at the subdistrict level. The outcome of this analysis is leaving school of that year. Y=1 indicates the individual left school in that year, and Y=0 indicates the individual did leave school in that year.
*** p<0.01, ** p<0.05, * p<0.1

Table 23: The Effect of Earthquakes on the Annual Hazard in Pregnancy

	All Individuals			**Urban**			**Rural**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake t= 0	-0.00809	-0.00782	0.00664	-0.0309***	-0.0298**	-0.0174	0.0193	0.0191	0.0371
	(0.0118)	(0.0120)	(0.0245)	(0.0114)	(0.0117)	(0.0106)	(0.0388)	(0.0389)	(0.0520)
Earthquake t= -1	-0.00721	-0.00734	0.00703	-0.0290**	-0.0286**	-0.0167	0.0131	0.0128	0.0351
	(0.0116)	(0.0116)	(0.0241)	(0.0114)	(0.0115)	(0.0111)	(0.0391)	(0.0392)	(0.0463)
Observations	40,005	39,184	15,244	8,878	8,719	6,131	13,921	13,718	9,113
R-squared	0.027	0.027	0.031	0.027	0.028	0.029	0.037	0.039	0.038
Number of BPS	1,398	1,394	524	343	340	277	464	457	328
Age	X	X	X	X	X	X	X	X	X
Personal	-	X	X	-	X	X	-	X	X
Household	-	-	X	-	-	X	-	-	X
Mean of Y	0.0165	0.0159	0.0143	0.0132	0.0126	0.00946	0.0228	0.0229	0.0176

All regressions control for fixed effect and personal variables, and standard error is clustered at the subdistrict level. The outcome of this analysis is pregnancy of that year. Y=1 indicates the individual is pregnant, and Y=0 indicates the individual is not pregnant.
*** p<0.01, ** p<0.05, * p<0.1

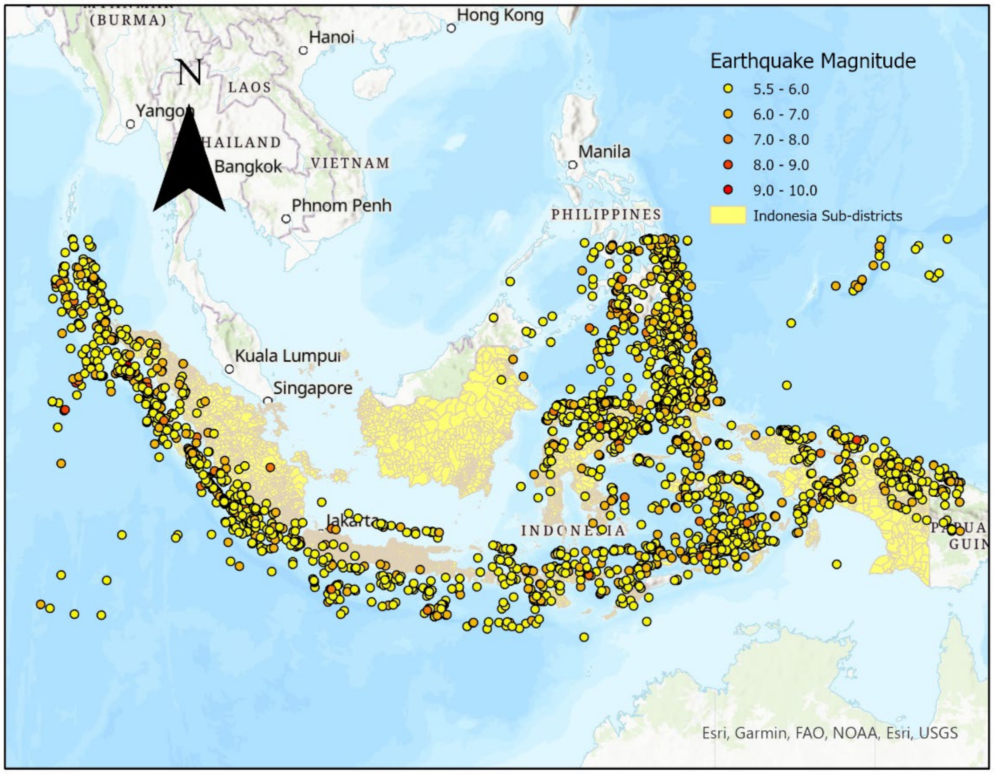

Figure 7: Earthquakes in Indonesia (1980-2015)

Note: The map shows the distribution of earthquakes from 1980 to 2015 with a magnitude higher or equal to 5.5 in Indonesia.

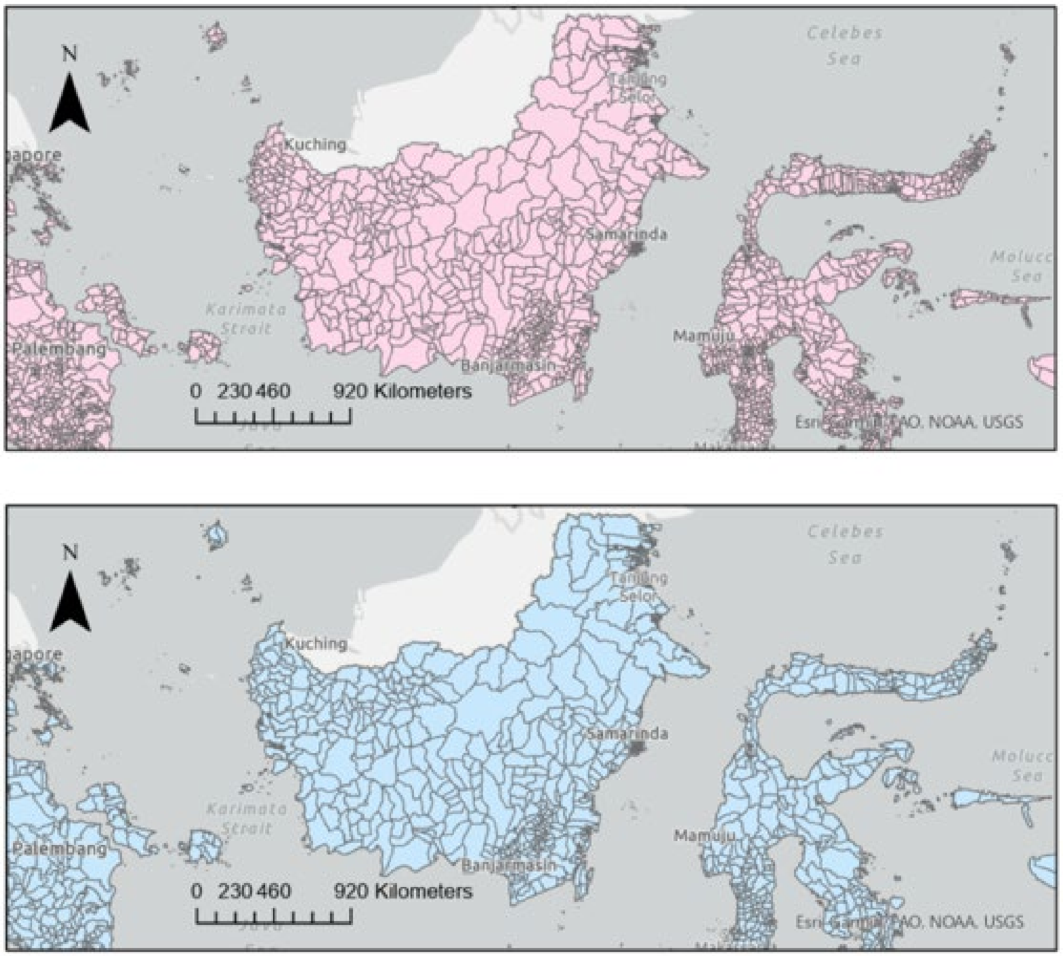

Figure 8: Subdistrict Before and After Merging using the Crosswalk

Note: This map only includes a part of Indonesia; the cut version shows how merging was performed. The map on the top indicates the subdistrict before merging, and the bottom one shows the subdistrict after merging. Before the merge, there were 6,695 subdistricts recorded by GDAM. After the merge, there are 3,881 subdistricts recorded by this study.

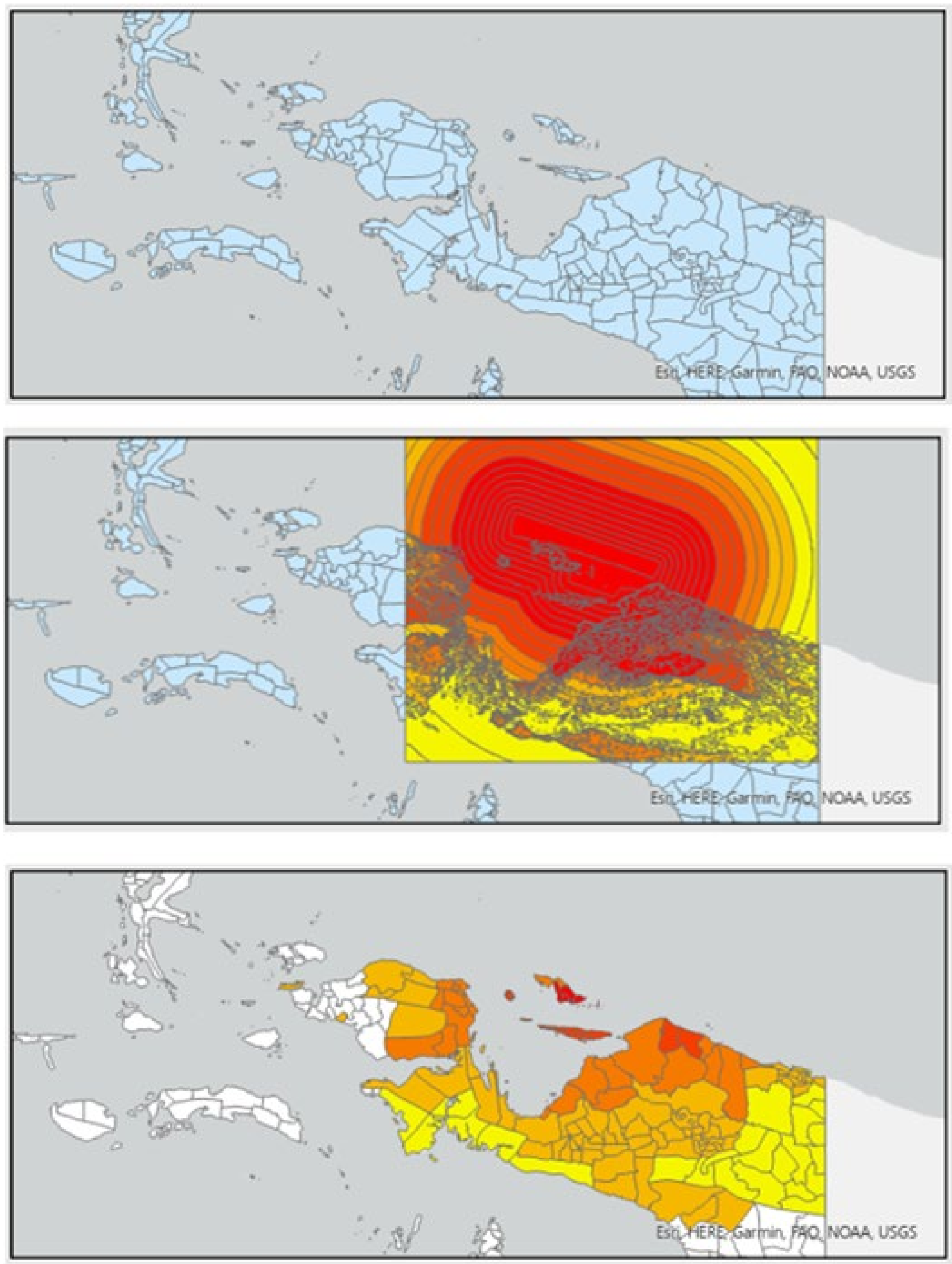

Figure 9:An Example of Merging Subdistrict and Earthquake Shapefile

Note: Those three maps include how connecting the subdistrict and earthquake shapefiles was performed. The final shapefile contains the average of the earthquake MMI at each subdistrict level. Darker red indicates a higher MMI level, and white indicates no earthquake shapefile overlaps within that subdistrict.

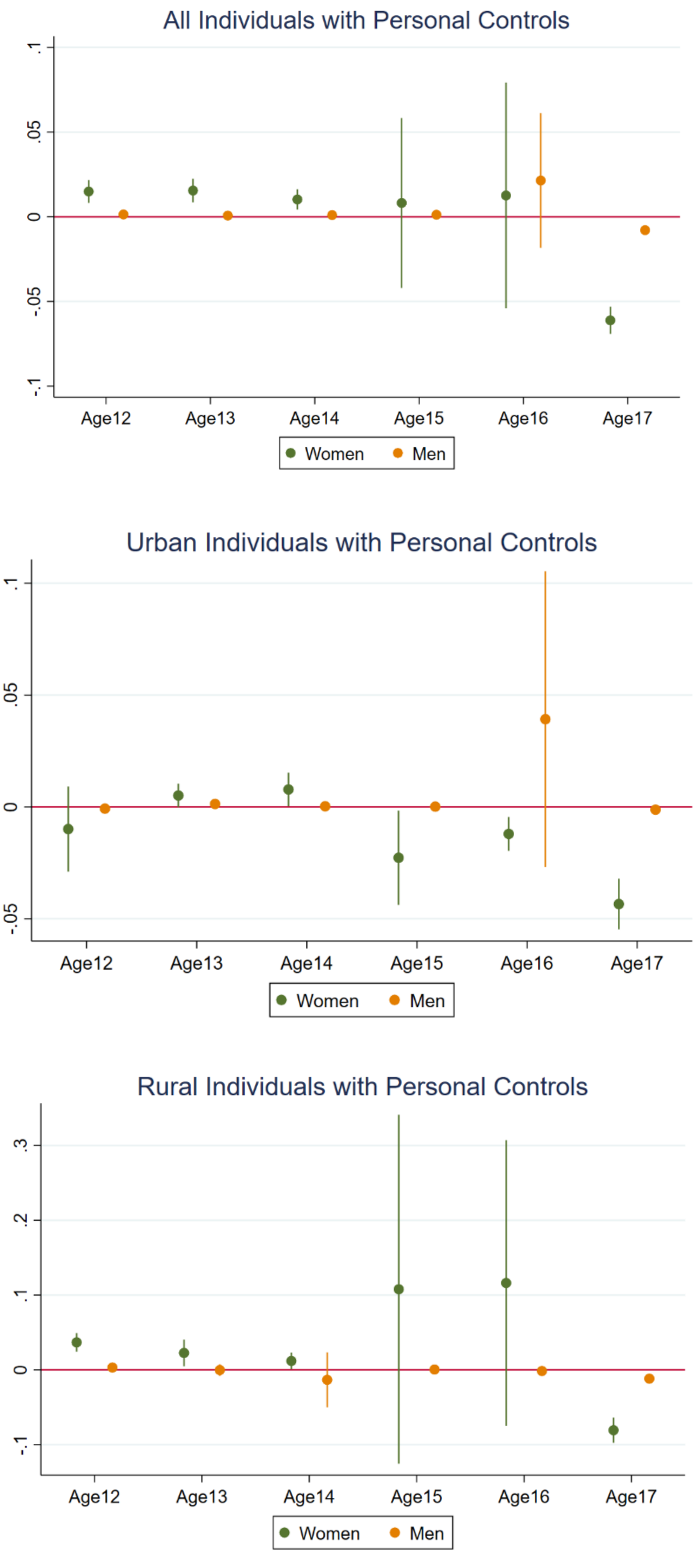

Figure 10: Impact of Earthquakes on Child Marriage by Ages

CHAPTER THREE: NATURAL DISASTERS, EDUCATION, AND ECONOMICS DEVELOPMENT

Over the past years, with the advancement in the GIS technology and improvement in the empirical strategies, there is emerging research on studying the effect of natural disasters on education attainments (Andrabi et al., 2021; Di Pietro, 2017; Nordstrom & Cotton, 2020; Sulistyaningrum, 2017; Thamtanajit, 2020). The effect of natural disasters on education could come from different channels, direct and indirect. There is a direct effect of natural disasters on education, where the disruptive power of natural disasters directly damages and disrupts the local communities, negatively impacting education and the probability of continuing the education process in the affected area (Kousky, 2012). Examples are damage to the physical capital, loss of educational resources such as teachers and schools, a rise in disability, and loss of human lives (Marcotte & Hemelt, 2008; de la Fuente & Fuentes-Nieva, 2010). On the other hand, there is an indirect effect. For example, the negative income shock from natural disasters may force individuals to leave school and pursue economic activities for additional income (Gitter & Barham, 2007). In addition, depression and other psychological effects might also indirectly result in education loss after natural disasters (Mamun et al., 2019; Kemp et al., 2011; Tian et al., 2022). Together with all effects, the existing research observes that natural disasters significantly disrupt the education process and result in human capital loss. I use the Yogyakarta earthquake as a natural experiment to study the effect of the earthquake on education attainments and explore possible mechanisms behind the impact of natural disasters on education attainments.

On 27 May 2006, at 05:53:58 am, the Yogyakarta earthquake struck Java Island in Indonesia with a magnitude of 6.3 on the Richter scale and a shallow depth of 10 m below sea level (Java Reconstruction Fund, 2009; Kirchberger, 2017). Although the magnitude of the earthquake was not deadly as the 2004 Indian Ocean Tsunami, it was among the costliest natural disasters in developing countries and resulted in US$ 3.1 billion in damage and losses (Elnashai, Kim, Yun, & Sidarta, 2007). The earthquake killed over 5,700 people, and about 280,000 houses were destroyed (Java Reconstruction Fund, 2007). As this earthquake largely disrupted the local communities and damaged the facilities, it provides a natural experiment on how significant disruptive events affect education attainments in developing countries.

To provide causal estimates of the effect of the earthquake on education, I utilize a difference-in-differences approach. My identification strategy exploits variation along two dimensions: compare how exposure to earthquake affected education attainment in the same birth cohorts across different subdistricts and compare the same subdistrict at-risk individuals who were younger than 18 at the time of the earthquake with individuals who were at least 18 years old, within the same subdistrict. My result suggests that being exposed to the earthquake, the years of education acquired decrease by about 1.2-2.6 years or a 10% to 20% reduction in the sample average. This effect is more pronounced among boys than girls, but the difference is not statistically significant at the conventional level. No significant effect on the probability of finishing compulsory school, but the probability of an individual finishing decreased by 6-10% percentage points, counting for a 6-10% decrease in the sample average. By comparing individuals living in the communities affected by the earthquake, the probability of boys being employed increased by 33.7 percentage points, and the probability of individuals younger than 18 working in the construction industry increased by 1.09 percentage points, with a more substantial effect on boys. Both results show evidence that boys might have left school to work, reducing education and disrupting their human capital accumulation process. To further explore the effect of the reduced human capital in the affected area, I track the nighttime light change before and after the earthquake at the community level. I find that the economic activities of the affected communities grew faster after the earthquake in the short term, not the long term. The reconstruction program will most likely explain this temporary growth change. My findings indicate that the earthquake significantly disrupted the human capital accumulation process in the affected area, which could explain the long-term economic development drop.

I contribute to the economics literature from two perspectives. First, I add to t the current literature on the causal effect of natural disasters on education. The impact of natural disasters on education has reached a conclusion where an adverse effect among affected individuals is observed in the short term and long term, but the result varies depending on the geographic locations (De Vreyer et al., 2014; Drabo & Mbaye, 2011; Caruso & Miller, 2015; Caruso, 2017). The current literature has already explored a few mechanisms of how natural disasters could affect education attainments. Being credit constrained and suffering from negative income could push individuals to leave school for work (Gitter & Barham, 2007). Depression and other psychological effects might also indirectly result in education loss after natural disasters from both the individual and household levels (Mamun et al., 2019; Kemp et al., 2011; Tian et al., 2022). Forced early marriage due to consumption smoothing and seeking safety after natural disasters might reduce education attainment (Alston et al., 2014; Corno Hildebrandt & Voena, 2020). Damage of the physical capital, loss of educational resources such as teachers and schools, a rise in disability, and loss of human lives might also contribute to education loss (Marcotte & Hemelt, 2008; de la Fuente & Fuentes-Nieva, 2010). A concurrent paper by Dong &

Yang (2020) uses the same event and datasets to study the effect of natural disasters on children's school and work activities. My paper is different from their paper in the following factors: I use cohort study design to more accurately identify the effect of the earthquake on the affected children. I track the migration history of the individual to more accurately the individual residential location at the time of the earthquake. I expand the outcome of interests in education attainments rather than just looking at total years of education and propose and verify a potential mechanism to the observed results. Thus, I add to the current literature by ascertaining the effect of natural disasters on education attainment and proposing and verifying a channel in which higher wages from the reconstruction programs may provide the incentive to individuals leaving school for short-term high paid working opportunities but suffering from long-term education loss and human capital accumulation.

This paper also closely relates to the existing literature on the macroeconomic effects of natural disasters. Most current literature finds that natural disasters have negative or inconclusive impacts in developing countries. Natural disasters destroy production capital, disrupt economic activities, and lower expenditure, eventually reducing the aggregate GDP in the short term (Anttila-Hughes and Hsiang, 2013; Bergholt and Lujala, 2012; Cavallo et al., 2013; Strobl, 2012). The negative effect of natural disasters is more potent in developing countries due to the difference in the credit market, governments policies, and after-disaster management, and lack of safety net (Guarnacci, 2012; Linnerooth-Bayer et al., 2007; Panwar and Sen, 2018; Sardar et al., 2016; Zhang and Managi, 2020). However, the long-term effect of natural disasters is still ambiguous. Some studies also found that the growth will likely bounce back to the before-disaster growth path in the long run (Strobl, 2012), but some countries may experience permanent growth reduction (Hsiang and Jina, 2014). By contrast, emerging literature also indicates that natural disasters could generate positive economic growth in the affected area afterward due to extensive after-disaster reconstruction programs, technology innovation, and replacement of production capital (Chen et al., 2021; Fomby et al., 2011; Hallegatte and Dumas, 2009; Heger and Neumayer, 2019). My result suggests that natural disasters related reconstruction could fuel economic growth after the earthquake when extensive reconstruction programs are implemented, but this effect could be short-term. Reduced human capital from natural disasters could reduce long-term economic development.

For the rest of this paper, section 2 presents the data resources and summary statistics, section 3 demonstrates the empirical strategy, Section 4 discusses the results, and the last section concludes the paper.

Data

I use the Indonesian Family Life Survey (IFLS) to obtain personal and household information to obtain education outcomes and other control variables. Secondly, by tracking the individual level migration history, I identify the individual residential location at the time of the earthquake at the subdistrict (Kecamentan) level, the smallest publicly available administrative unit tracked by IFLS. After locating the residential location, I integrate the personal and household-level information with the GADM dataset. This approach allows me to identify the subdistrict GPS location and assign each individual with subdistrict level information based on their migration history. Lastly, using the GPS location of each subdistrict, I find the earthquake intensity based on United States Geological Survey (USGS) ShakeMap data to identify the overall impact of the earthquake recorded by Modified Mercalli Intensity (MMI). Lastly, I use National Oceanic and Atmospheric Administration (NOAA) for the nighttime light data to track the economic activities change at the community level. This section demonstrates the data and data cleaning process in detail.

Indonesian Family Life Survey (IFLS)

Indonesian Family Life Survey (IFLS) is an ongoing longitudinal survey in Indonesia asking the respondents of their individual, household, and community-level information. The sample represents 83% of the Indonesian population and contains over 30,000 individuals living in 13 of the 27 provinces in the country (Strauss, Witoelar, and Sikoki, 2016). The first wave was in 1993, the second wave was in 1997, the third wave was in 2000, the fourth wave was in 2007/2008, and the last wave was in 2014/15. I use IFLS5 as the main survey to extract personal information. Due to the complex nature of IFLS, some panel information is extracted from the earlier wave if the latest wave information is missing. The main outcome of interest is the years of education, the probability of finishing compulsory school, the probability of finishing high school, and the probability of dropping out of high school. A detailed definition of those variables is presented in Table 1.

My analysis includes a vector of personal and household characteristics to count for unobservable factors that could have affected the outcome of interests. I include ethnicity, religion, locality by the time of the earthquake, and birth year for personal level variables. I also have a vector of household characteristics variables as control variables. However, variables like socioeconomic factors are very likely to be impacted by the earthquake. To ensure the strict exogeneity of those variables and isolate the direct impact of the earthquake, I use household information collected in 2000 from IFLS3, the wave before the earthquake. After obtaining the household characteristics variables in IFLS3, I link the individual with household characteristics from IFLS based on their cross-wave personal IDs. I include the number of children living within the family, the birth order of the individual within the household, the mother's education, an indicator recording if the household has higher than the average first component of the principle component analysis(PCA) of the household durable assets to count for socioeconomics condition (McKenzie, 2005). A detailed definition of those variables is presented in Table 1.

ShakeMap and GDAM

I link the U.S. Geological Survey (USGS) ShakeMap Data and Database of Global Administrative Areas (GADM) to estimate the effect of the earthquake at the subdistrict level within Java Island (Kecamentan). USGS ShakeMap records the earthquake intensity at a specific geographic unit, recording the actual effect and intensity of the earthquake by Mortified Mercalli Intensity (MMI) (USGS, 2021). Higher MMI indicates a higher level of ground shaking movement and provides a better proxy to estimate the effect of the earthquake than other more traditional measurements, such as distance to the epicenter, by reducing the geographic heterogeneity of the space and limiting other confounding factors in the actual effect of the Earthquake (Wald et al. 1999). GADM records 6695 subdistricts in an area of 1,877,519.0 km^2, and BPS (2010) reports a total population of 237,641,326 in Indonesia, giving the average population per subdistrict of 35,495 and an average area per subdistrict of 280 km^2 (BPS, 2010). With the information from both datasets, I use Arc-GIS Pro to find the average MMI within a subdistrict to identify the earthquake's intensity at the smallest administrative unit tracking the individual migration history and their residential location by the time of the earthquake using publicly available datasets.[22] In addition, since the ShakeMap only records intensity levels higher than 2.8 due to technical limitations, some areas do not have a recorded intensity. Therefore, I

[22] ArcGIS Pro summarizes the MMI intensity using the arithmetic average of all polygons within that subdistrict. If the border of the subdistrict crosses the polygons, ArcGIS Pro calculates the polygons within the border and assigns a weighted average to the calculation.

assume that regions outside the covered areas have MMI intensities as 0 since MMI 0 to 2.8 introduce minimal effect in the affected area (Wald et al. 1999).[23] Figure 1 shows the Modified Mercalli Intensity in Java Island at the subdistrict level. The darker blue indicates a higher Modified Mercalli Intensity, and the white color indicates no effect or intensity recorded by ShakeMap.

The earthquake's epicenter was near Yogyakarta city, one of Indonesia's biggest cities. Including individuals from rural or less underdeveloped areas might bias the earthquake's effect, leading to confounding effects in the results. Following Kirchberger (2017), I select subdistricts centered within 50 kilometers of the major cities buffer with 1-kilometer accuracy to construct a more homogenous sample and eliminate the heterogeneous effect due to the locations. Appendix Table A.1 lists all major cities and their GPS coordinates used in this study.[24] Figure 2 illustrates the subdistricts selected for this study by using ArcGIS Pro. The dark blue indicates that those subdistricts are located within the 50-kilometer buffer of major cities within Java Island. All other subdistricts or white subdistricts are excluded from my study.

Linking IFLS and the Earthquake

To assign the intensity to each individual, I take advantage of both GADM and IFLS to share Badan Pusat Statistik (BPS) codes at the subdistrict level. BPS code is the Indonesian government code system recoding the area codes at five administrative divisions: country, province (provinsi or daerah istimewa), district (kabupaten) or city (kota), subdistrict (kecamatan), and village (desa or kelurahan).[25] Because the subdistricts were split or merged into

---

[23] After an intensity of 3 the actual difference of MMI is minimal, and this level of intensity typically does not lead to any earthquake damage anymore (Wald et al. 1999).
[24] Those cities had more than 300,000 residents in Java based on the Atlas of Urban Expansion data in 2000.
[25] To protect privacy, IFLS only reports down to the subdistrict level information, not the village information. Therefore, the subdistrict is the smallest administrative unit can be studied.

new subdistricts or for other political reasons, many of the BPS codes changed during the study period. Thus, I utilize the crosswalk provided by IFLS to convert all BPS from IFLS5 to IFLS4 as IFLS4 fits the GADM area code better. Specifically, as some BPS code has changed from 2007/2008 to 2014/15, I change the new BPS code to the old version to better match the BPS code from GADM, a dataset published in 2010.[26] After converting the IFLS5 BPS to the IFLS4 BPS code using the crosswalk, I successfully matched all the GDAM BPS code subdistricts with the IFLS BPS code.[27]

To study the impact of the earthquake education outcome, one would like to measure earthquake intensity in the subdistrict of residence at the time of the earthquake. Unfortunately, the IFLS did not interview the respondents right before the earthquake. Using the location after the earthquake is not optimal because I expect individuals to move in and out of the affected locations after the earthquake. In this case, only using the survey time residential location would create endogeneity issues and bias my results. To correct this, I utilize the detailed IFLS migration history at the individual level to identify their residential location by the time of the earthquake. IFLS migration history records the residential location of the birthplace, and then it records the residential location at 12 years old if the individual moved. Then, it provides detailed information on migration locations in chronological order. Thus, I keep updating the individual's residential location since birth and stop updating at a cutoff time, which is the earthquake time in

[26] I convert the BPS code of the Java Island provinces since all other provinces are out of the scope of my study. There are two converting cases: 1) a subdistrict level information/name change and 2) larger subdistrict split into smaller ones. For subdistrict 3218010 in IFLS5, it changed name from Ciamis to Pangandaran and I covert the BPS code back to 3207010, which is the original code in IFLS4. For other subdistricts, they split from a larger subdistrict to smaller subdistricts from IFLS4 to IFLS5, which include subdistrict including BPS code of 3522191, 3575031, 3575031, 3217071, 3604251. I convert the IFLS5 code back to IFLS4 so that 3522191 to 3522190, 3575031 to 3575030, 3217071 to 3217070, 3604251 to 3604250.
[27] only some special geographic units from GDAM are dropped from my dataset. Such geographic units including reservoirs, forests, and dams which do not have residents.

May 2006.[28] For my study, since IFLS5 was conducted in 2014/2015, individuals who are 18, 19, and 20 years old would have been less than 12 years old in 2006. Therefore, to eliminate the endogeneity issue of the survey information, I use their birthplace as the residential location rather than the residential location at 12 years old.

After obtaining the residential location or BPS at the time of the earthquake, I match each individual with earthquake data through the BPS code. Each subdistrict has a unique BPS and an associated intensity level, and each individual has a BPS code that records their residential location by the time of the earthquake.[29] By matching datasets together, I acquire individual-level information, precisely recording whether they were affected by the earthquake. I further define individuals exposed to the earthquake if they experience MMI higher than 6 at the subdistrict level (Wald et al. 1999). Table 1 summarizes the key variables used for this paper.

Empirical Strategy

To provide causal estimates of the effect of the earthquake effect on child marriage, I utilize a difference-in-differences approach. My identification strategy exploits variation along two dimensions. First, I use variations in earthquake intensity across geographic locations. Second, I rely on variation in exposure to the earthquake across birth cohorts (Duflo 2001; Caruso and Miller 2016). [30]

[28] For instance, if an individual moved in 2007, their residential location would not be updated or replaced. Only the previous residential location would be recorded to indicate the residential location at the time of the earthquake.
[29] Note that about 10 percent of the personal BPS codes have errors, in which no such BPS codes could be found in the actual database. Another 10 percent of the individual did not know their subdistrict code (BPS ended with 999) or the interviewers cannot identify such places in the data (BPS ended with 998). Therefore, I only successfully matched 80 percent of the individuals with an MMI intensity at the subdistrict level.
[30] Since I only have a single cross-section data from IFLS5, I cannot distinguish between age and year of birth and control for age and cohort fixed effect at the same time.

When comparing across space, I explore how exposure to the earthquake affected the educational outcomes in the same birth cohorts across different subdistricts. I define the individual affected by the earthquake or earthquake equals to 1 if the MMI is greater or equal to 6 or equal to 0 if the MMI is less than 6. When comparing across birth cohorts, compare the same subdistricts at-risk children younger than 18 at the time of the earthquake ("exposed") with individuals at least 18 years old within the same subdistrict. Since the average years of education are 10 to 11 years in the main sample, the critical age of ending education is about 16 to 18 years, assuming the education starting age is 6 to 7 years old. Thus, I assume individuals younger than or equal to 18 but older than 12 are the affected cohorts whose education process could have been affected by the earthquake. For individuals older than 18, although their college education or more advanced degrees could still be affected by the earthquake, they would not be the popular or major groups in the sample. Thus, I define cohorts older than 18 as unexposed cohorts whose education outcomes are unaffected by the earthquake. This definition is verified in the robustness check section by including ages 12 to 22 as affected cohorts to verify that the result is not affected by choice of ages. The difference in difference strategy can be generalized in the following regression format:

$$\mathrm{Y_{ist}} = \mathrm{\alpha_0} + \alpha_1(\mathrm{Earthquake}_s \times \mathrm{Affected\ Cohort_t}) + \gamma \mathrm{X_{ist}} + \mathrm{\delta Z_{ist}} + \mathrm{\theta_s} + \vartheta_{\mathrm{t}} + \mathrm{\varepsilon_{ist}}$$

$\mathrm{Y_{isc}}$ is a vector of education outcomes, including years of education, and the probability of finishing compulsory school or senior high school. The main outcome of interest is years of education. It is a continuous variable that equals the total years of education acquired if person *i* in subdistrict *s* and belongs to cohort *t*. $\mathrm{Earthquake_s}$ is the earthquake indicator equal to 1 if the affected subdistrict has an average MMI equal to or greater than 6 in subdistrict *s*. Affected $\mathrm{Cohort_t}$ equals 1 if the individual was born between 1995 and 1989 or 12 to 18 at the time of the earthquake and 0 if they were born between 1988 and 1982. I exclude individuals who were born outside of this time frame. The difference-in-differences coefficient α1 records the causal impact of the earthquake on outcomes of interests.

X_{isc} is a vector of individual characteristics, including gender, ethnicity, religion, and residential type at the time of the earthquake. Z_{isc} is a vector of household characteristics, including the mother's education, the number of children, birth order, and indicator if the household has the first component of PCA in the durable assets higher than the median level in IFLS3. As the earthquake could affect wealth and education in the household, all household characteristics are measured before the earthquake, so they are predetermined with respect to the earthquake. Ideally, one should also select household characteristics before education decisions were made, but this process would require assigning earlier household characteristics and accessing the time of leaving school, which is out of the scope of this study. Additionally, subdistrict fixed effects θ_s are included to capture all unobserved time-invariant characteristics common within a subdistrict at the time of the earthquake, and cohort fixed effects ϑ_c are included to capture all unobserved factors that are common among birth cohorts. The standard errors are clustered at the subdistrict level to allow for serial correlation within the same geographic area.[31]

The identifying assumption of the difference-in-differences strategy is that, in the absence of the earthquake, trends in the outcome variables across cohorts of birth would have been the same in subdistricts affected by the earthquake and subdistricts not affected by the earthquake. Although this parallel assumption is inherently untestable, I provide indirect support for the

[31] There are 679 unique subdistricts in this study and about 7.09 observations per subdistrict.

analysis using a falsification test by the end of this paper. The falsification test supports the assumption and provides evidence for the validity of the study design.

Main Results

The main interest of the outcome is years of education. Table 2 summarizes the results by three groups: all, female, and male. In general, exposure to earthquakes reduces the total years of education by 0.5 to 0.7 years, depending on the specifications, but those results are not precisely estimated at the conventional statistical level.[32] Thus, a subgroup analysis reveals that males face a reduction of 1.2 to 2.6 years of education, reducing the affected individual years of education by 11.1 % to 23.6% compared to the sample average. This effect is estimated precisely. For females, the effect of the earthquake is ambiguous and cannot be concluded. This result indicates that the years of education acquired decreased among individuals affected by the earthquake during their critical human capital accumulation period. This effect is more pronounced among boys than girls, but the difference is not statistically significant at the conventional level.

Brown and Velásquez (2017) indicate that although years of education might show human capital accumulation, the benchmark or the school level is a more meaningful measurement for education attainments in developing countries. To better understand how being exposed to earthquakes might have affected the probability of finishing certain school levels, I split the education into two new variables: 1) the probability of finishing compulsory school, which is defined as individuals finished junior high school, 2) the probability of finishing senior high school, which is defined as individuals finished the senior high school after they attempted. For the probability of finishing senior high school, I restrict the sample to those who attempted

[32] The Chow-Test does not indicate heterogenous effect between male and female at conventional level.

high school better to understand the effect of the earthquake on individual behavior. The Indonesian school system has very different types of schooling, including general school, religious school, adult education, and vocational school. Therefore, I define those variables regardless of which school system they were in.

Table 3 presents how the earthquake affected the probability of finishing compulsory school among affected cohorts compared to unaffected cohorts. The results show that exposure to the earthquake during school age could either decrease or increase the probability of finishing junior high school or compulsory school and all coefficients are not precisely estimated. As the age to complete compulsory school is aging from 14 to 16, I defined the affected cohort accordingly based on their critical time of finishing a compulsory school by limiting to cohorts between the ages 14 to 16 and comparing them with the age of 17 to 19. The result is similar to the specification used in the main analysis, with no observed effect.

To further delve into the effect of the earthquake on education after compulsory school, I restrict the sample to those who went to senior high school and create another indicator equal to 1 if an individual finished high school after attempting. This indicator records if an individual attempted high school and completed it. Table 4 shows the results of the earthquake on this outcome. The probability of attempting and finishing high school decreased by 6 to 10.6 percentage points for the whole sample, depending on specifications. This effect is larger and more significant among males than females but not statistically different. Although the statistical power was reduced after splitting males and females into two groups, it is likely to be explained by the fact that the number of observations was reduced. But the effect on males is still estimated precisely at 0.05 level. This result indicates that exposure to the earthquake during school age reduces the probability of completing senior high school among those who attended high school.

Mechanism

Kirchberger (2017) indicates that the earthquake significantly affected the local labor market, where the reconstruction program attracted a large inflow of labor moving from the agriculture industry to the construction industry due to higher wage levels. Her result indicates that the boys are possibly attracted by the high wages of the construction projects, which forces them to leave school and join the labor market. To verify this hypothesis, I extract data from the IFLS3 and IFLS4 to compare the probability of an individual employed at the time of the survey and the probability of an individual working in the construction industry. As IFLS only report individual work history if they are out of school or have worked before, the sample for this analysis is individuals who were not in school or had worked during the survey time. I use IFLS4 instead of IFLS5 for this analysis to count for the instant effect of the earthquake on the labor market. As IFLS4 was collected in 2007/2008, it shows the immediate effect of the earthquake and confirms if the reconstruction project has attracted more boys into the local market, especially for the construction industry. To isolate the effect of the earthquake on young individuals, I limit all samples to the age between 15 to 18, where 15 is the youngest age in the IFLS survey collecting for working condition information and 18 is the critical age used in the main specification of this study. Specifically, I follow the specification as follow:

$$\mathrm{Y_{ict}} = \beta_0 + \beta_1(\mathrm{Earthquake}_c \times \mathrm{IFLS4}_t) + age + \theta_\mathrm{c} + \vartheta_\mathrm{t} + \varepsilon_\mathrm{ict}$$

Where Y_{ict} is the outcome of interests stands for the probability of an individual being employed and the probability of an individual working in the construction industry at the time of the survey. It equals 1 if individual *i* living in an IFLS community *c* during the IFLS survey *t*. Earthquake_c equals 1 if the community has earthquake MMI intensity higher than 6, or 0 otherwise.[33] IFLS4_t equals 1 if the respondent is from IFLS4, or 0 otherwise. β_1 presents the causal effect of the earthquake on the outcome of interests. Age is the age as a continuous measurement, θc is the community fixed effect, and ϑt is the survey wave fixed effect, all counting those unobservable characteristics that could affect the outcome of interests across age, time, and location. A similar procedure is done for IFLS 5 sample to explore if this effect could be sustained in the long run. The standard errors are clustered at the community level for serial correlation. Table 6 presents the results.

The result indicates for the sample who are between the age of 15 to 18 and who are not in school and worked before, the probability of they are employed increased by 7.8 percentage points, increased by 33.7 percentage points for males, and 0.4 percentage points for females compared to those who live in communities not affected by the earthquake after the earthquake. The estimation is precisely estimated for males at 0.05 level. The Chow test indicates that the difference between males and females is not significant at the conventional level. The probability of occupation in the construction industry increased for males and females, counting for about a 142% increase in the sample average, where most of the power resides within boys, as the probability of working in the construction industry increased by 164% compared to the sample average. Although the estimation indicates the probability of females working and working in construction also increased, they are not precisely estimated. This set of results implies that boys might have left school for work and started working in the construction industry, paying higher

[33] For simplicity, I use the community level information, which is a geographic unit recorded by IFLS. The subdistrict level information is only available at the household level, and the BPS code could change over time.

wages and salaries than other jobs. This behavior might explain why the earthquake disrupted their education attainment among boys, not girls, as shown in the main results.

When compared to IFLS5, the probability of individuals who are between the age of 15 to 18 and out of school at the time of the survey is not different between communities affected by the earthquake or not, but the probability of them working in construction is still higher among communities affected by the earthquake. This result indicates that the effect of the earthquake on the local labor market is likely to be temporary but permanently change the economic structure, as Kirchberger (2017) indicated.

Nighttime Light and Economic Development

To further explore if the earthquake changed the economic development within the affected communities and explore the effect of individuals moving from school to the labor market, I use the community fixed effect model to estimate the causal impact of the earthquake on the nighttime light change at the community level by following Heger and Neumayer (2019). Community is a unit created by IFLS, and no information records the community-level GDP. To study the long-term effect and measure economic activity change, I use nighttime light data from the National Oceanic and Atmospheric Administrative (NOAA) to approximate the economic activities. By studying the change of the nighttime light, I can identify the earthquake's impact on economic development within the affected area (Fabian et al., 2019; Heger & Neumayer, 2019). NOAA provides maps of annual average nightlights where the highest resolution is a pixel of size about 30 × 30 arc-seconds. The lights are recorded based on the brightness in Digital Numbers (DN) ranging from 0 to 63, where 0 indicates no detected light and 63 indicates the highest level of luminosity by excluding natural glare and moonlight. Economic literature often uses nighttime light to measure economic activities. A larger DN change suggests a higher level of economic growth, and a lower change or negative change of DN indicates a lower-level development or economic recession (Heger & Neumayer, 2019). Therefore, I extract nightlight data from 2000 to 2013 to examine the DN changes in the affected area compared to the non-affected areas, six years before the earthquake and six years after the earthquake. Three satellites recorded the satellite image from 2000 to 2013: F15 records images from 2000 to 2007, F16 records images from 2008 to 2009, F17 records images from 2010 to 2013 (NOAA, 2021). Although F16 also provides images from 2005 to 2007, I use F15 until 2007 to decrease the measurement error due to the satellite detection difference. Figure 3 demonstrates the change of the nighttime light at Java Island in different years.

I extrapolate the DN change within a community area to measure the changes in economic activities. Considering nighttime light suffers from calibration issues that do not accurately identify the exact location on the map, using the summation of luminosity within an area rather than a specific location helps to reduce the measurement error and better estimate the economic changes within that region (Gibson et al., 2021). Specifically, I spatially link the community GPS location with nighttime light raster data, giving a sum of DN for each community within a 10 kilometers buffer from 2000 to 2013. In other words, I create a buffer of 10 kilometers around all selected communities to measure the luminosity changes within this buffer. This step helps to ensure that I measure the total change of the DN within the community area rather than just a GPS point. Moreover,

To estimate the change in nighttime light, I estimate the following specification.

$$\Delta NL_{ct} = \beta_0 + \beta_1 \sum_{t=2001}^{2004} MMI_c \times T_t + \beta_2 \sum_{t=2006}^{2006} MMI_c \times T_t + \beta_3 \sum_{t=2007}^{2009} MMI_c \times T_t$$
$$+ \beta_4 \sum_{t=2010}^{2013} MMI_c \times T_t + \sigma_c + \rho_t + \varepsilon_{ct}$$

In the estimation equation, ΔNL_{ct} is the change of the community level nighttime light from the previous year, and *NLct* is the log transformation of the summation of the DN of pixel *n* on the map within 10 kilometers of the community *c* in year *t.*[34]

$$NL_{ct} = \log\left(\sum_{i=i}^{N}(DN_{nct}\right)$$

Thus, β_1 records the nighttime light change between the affected and non-affected communities from 2000 to 2004. I dropped the change between the year 2004 to 2005 as the comparison group. If the communities had a similar economic growth rate, the $\beta 1$ would be statistically insignificant between those with high MMI intensity and those with low MMI intensity.[35] β_2 captures the direct damaging impact of the earthquake in 2006 compared to the change from the year 2004 to 2005. If the earthquake significantly impacted and destroyed numerous facilities in the affected area, $\beta 2$ would be negative due to the earthquake damage. The aid and reconstruction period is from 2007 to 2009, recorded by the coefficient of β_3. The sign of β_3 demonstrates the effect of the reconstruction and the earthquake: a negative sign indicates the positive impact from reconstruction is less than the negative effect from the earthquake, and a positive coefficient indicates the reconstruction provides a faster growth rate in those areas

[34] Heger and Neumayer (2019) added 0.01 DN to each pixel point to avoid 0 DL points in the log transformation.
[35] As I measure the change of nightlight in the affected area, 2000 is dropped from the analysis as it is the start year for the regression.

with the high earthquake intensity. The long-term economic growth difference is recorded by β_4, in which a positive sign indicates economic growth sustained in the affected communities, and a negative sign indicates economic growth collapsed in the affected communities after reconstruction. σ_c is the community fixed effect, and ρ_t is the year fixed effect recording the unobservable heterogeneity due to the community and year difference.[36] The standard errors are clustered at the community level to allow serial correlation within communities.

Table 7 summarizes the results by dividing the time into four categories: before the earthquake (2001 to 2004), during the earthquake (2006), after the earthquake but during the reconstruction period (2007-2009), and after the earthquake and after reconstruction (2010 to 2013), in which the change between 2004 to 2005 is dropped as the comparison group. [37] From 2001 to 2004, when compared to the nighttime light change from 2004 to 2005, the economic growth of communities in the earthquake-affected area was slower than the unaffected communities by 2.15 percentage points. This result indicates that the nightlight change within affected communities was significantly lower than in the unaffected communities before the earthquake. Unlike the theoretical prediction, the affected communities did not suffer a considerable reduction in nightlight change from 2005 to 2006 during the year of the earthquake. On the contrary, the nighttime grew faster within the affected communities by 1.29 percentage points. During the reconstruction period, the affected communities had about 1.40 percentage points more nighttime light growth than the unaffected communities. It is about 58.6% above the average growth rate from 2000 to 2013 among all studied communities. Lastly, after the

---

[36] Three satellites record the satellite image from 2000 to 2013, F15 records images from 2000 to 2007, F16 records images from 2008 to 2009, F17 records images from 2010 to 2013. F16 shows clearly glares in the map, in which the darkest area shows some level of light all year around. Therefore, I choose F15 whenever it is possible.
[37] As I record the change of the nighttime light, year 2000 is dropped due to no comparison between 1999 to 2000. Moreover, year 2001 is the base year for comparison, which is dropped from the regression.

reconstruction was finished in 2009, the nighttime light growth rate was 2.39 percentage points lower in the affected communities compared to the unaffected communities. As I use the change of DN to evaluate the economic activity change, the change of DN would not have reflected the economic growth if a community had reached the maximum luminosity at DN of 63. Therefore, the second column uses the change of median rather than the change of the total summation of DN. The result is robust across two measurements, giving the exception that no significant difference is observed before the earthquake.

The results indicate that communities affected by the earthquake had slower economic development before the earthquake. Although the reconstruction period accelerated development, the effect diminished and returned the development track to the original level. Thus, the reconstruction and the influx of funds did not generate long-term economic development effects in the affected region. By linking economic development and the previous result, it could be true that boys leaving school and participating in the construction program fueled economic growth. However, this result also indicates that those opportunities are short-term. Once the reconstruction program is finished, they cannot seek such opportunities anymore, which may lead to a long-term adverse effect on their lifetime earnings. However, I cannot identify the causal relationship between economic growth and educational outcomes. On the one hand, it could be true that fueled economy attracted the boys to leave school and participate in the reconstruction program, leaving all of their education outcomes to decrease. On the other hand, it could be true that due to school closure and other disruptive events, boys left schools and participated in the labor market, resulting in a fueled economy. A more advanced study is required to find a definitive answer to this causal relationship.

Threats to Identification Strategy

The major threat to my identification strategy is that some unobserved subdistrict level preexisting trends could affect educational attainment. If true, those factors could also explain my results rather than attributing all estimated effects to the earthquake. To test if the effect estimated is biased by unobserved subdistrict-level preexisting trends, I follow Brown and Velásquez (2017) and use a falsification test to examine the validity of my empirical strategy. Specifically, instead of comparing the group at risk of school age between 12 to 18, I increase the at-risk age by six years. Thus, my exposed group is aged 19 to 25, and the non-exposed group is aged 26 to 32. As most individuals finish their education after age of 18 in Indonesia, the earthquake would not have affected their education decision, and the two age groups should not have different educational outcomes. Table 8 provides the results of this falsification test. The test supports the validity of my empirical strategy and estimated results from Table 2 and indicates no preexisting years of education before the earthquake.

Robustness Check

This section implements a few sets of robustness checks to check that the study results are not subject to choosing how I defined treatment and cohort. First, instead of defining the affected cohorts between the age of 12 to 18, I extend the age by four years to 22, allowing individuals to finish college-level education. Second, I define an earthquake as an MMI greater or equal to 5 instead of 6. Tables 9 and 10 summarize the results for years of education. Other outcomes results are similar to the main results but not reported in this section. The tables show that the estimated effect is similar to the main results, and the choices on how I defined affected cohort and earthquake do not change the results.

Conclusion

The existing literature agrees that natural disasters could affect human capital accumulation by disrupting the education processes. I estimate the effect of the 2006 Yogyakarta Earthquake on a vector of education attainments by comparing individuals affected and not affected by the earthquake before and after. The difference-in-difference study design provides evidence of the negative impacts of the earthquake on the total years of education and the probability of dropping out of high school, where those effects are more pronounced among boys than girls. This set of results indicates that the earthquake has disrupted the education process and imposed a larger negative effect among boys than girls. To delve into the mechanism behind the observed effect, I compare the probability of individual employment and occupation change among individuals aged 15 to 18. The results indicate that boys younger than 18 are more likely to be employed and work in the construction industry after the earthquake. This set of results implies that boys might have left school for work after the earthquake, disrupting their educational attainment. The result shed light on the after-earthquake management and construction. While the construction project is necessary, it might negatively affect affected individuals as they could leave school to seek relatively higher-paid jobs and never be able to return to school for human capital accumulation. Such short-term benefits but long-term negative impacts would result in a lasting negative effect among boys. In this case, providing a return to school or an opportunity to be educated again after natural disasters might be necessary.

To further explore the effect of the reduced human capital in the affected area, I track the nighttime light change before and after the earthquake at the community level. I find that the economic activities of the affected communities grew faster after the earthquake during the reconstruction period but not after the reconstruction program was finished. This temporary growth change is most likely to be caused by the reconstruction program and an influx of resources, as well as more school-age labor moved into the labor market at the time of the earthquake. However, as the effect diminishes in the long term, reduced human capital accumulation might also affect macroeconomic development. Combining all findings, the earthquake significantly disrupted the human capital accumulation process in the affected area, pushing more boys into the labor market. But the reduced human capital and education could explain the long-run economic development drop after the reconstruction, and temporary work opportunities ceased.

Tables & Figures

Table 24: Summary Statistics

Variables and Definition	Observations	Mean	Std. Dev.
Earthquake: =1 if MMI ≥ 6	4,816	0.0768	0.2663
Exposed: =1 if the birth year is between the Year 1989 and 1996	4,816	0.4400	0.4964
Years of Education: total years of education received	4,653	10.6884	3.2635
Compulsory School: =1 if the individual finished at least junior high school	4,664	0.8107	0.3918
Failed High School: =1 if the individual started high school but did not finish	4,664	0.0489	0.2157
Javanese: =1 if the induvial has Javanese ethnical background	4,816	0.6321	0.4823
Islam: =1 if the induvial is Muslim	4,816	0.9693	0.1726
Age: In years	4,816	26.2994	4.6856
Male: =1 if the individual is male	4,816	0.4616	0.4986
Village: =1 if the individual lives in a village by the time of the earthquake	4,798	0.5392	0.4985
Small Town: =1 if the individual lives in a small town by the time of the earthquake	4,798	0.2645	0.4411
Large Town: =1 if the individual lives in a large town by the time of the earthquake	4,798	0.1963	0.3973
Number of Children within the family	3,000	2.8567	1.4497
Birth Order Within Family: 1 indicates the first birth	3,000	1.9630	1.1576
Mother Primary: =1 if mother received less than primary level education	2,834	0.3073	0.4615
Mother Secondary: =1 if mother received less than secondary level education	2,834	0.5155	0.4998
Mother More Than Secondary: =1 if mother received more than secondary level education	2,834	0.1634	0.3698
Household Asset Median: =1 if the household has an asset first PCA greater than average	3,000	0.5743	0.4945

Table 25: The Earthquake Effect on Years of Education

Years of Education	All			Female			Male		
	(1)	(2)	(3)	(7)	(8)	(9)	(10)	(11)	(12)
Earthquake X Exposed	-0.767	-0.759	-0.567	-0.336	-0.318	1.360	-1.264*	-1.343	-2.636***
	(0.505)	(0.582)	(1.564)	(0.935)	(1.020)	(2.897)	(0.758)	(0.860)	(0.832)
Observations	4,653	4,178	2,431	2,526	2,261	1,238	2,127	1,917	1,193
R-squared	0.398	0.430	0.487	0.495	0.526	0.597	0.399	0.433	0.471
Personal Control	-	X	X	-	X	X	-	X	X
Household Characteristics	-	-	X	-	-	X	-	-	X
Fixed Effects	X	X	X	X	X	X	X	X	X
Mean of Y	10.69	10.78	11.09	10.64	10.74	11.11	10.74	10.83	11.07

Years of education is calculated by adding all years of education level acquired by the induvial, assuming that jumping or staying in a grade does not happen. Thus, elementary school takes 6 years, junior or senior school takes three years, college takes 4 years, a master's requires 2 years, and a doctorate takes 4 years. All specifications include subdistrict fixed effect and cohort fixed effect, and standard error is at the subdistrict level.

*** p<0.01, ** p<0.05, * p<0.1

Table 26: The Earthquake Effect on the Probability of Finishing Compulsory School

=1 If Finished Compulsory School	All			Female			Male		
	(1)	(2)	(3)	(7)	(8)	(9)	(10)	(11)	(12)
Earthquake X Exposed	-0.0095	-0.0140	0.0152	-0.0355	-0.0451	0.116	0.0154	0.0222	-0.0496
	(0.0492)	(0.0507)	(0.107)	(0.0749)	(0.0765)	(0.195)	(0.0784)	(0.0882)	(0.0719)
Observations	4,664	4,180	2,432	2,530	2,261	1,238	2,134	1,919	1,194
R-squared	0.342	0.395	0.412	0.431	0.491	0.517	0.353	0.394	0.401
Personal Control	-	X	X	-	X	X	-	X	X
Household Characteristics	-	-	X	-	-	X	-	-	X
Fixed Effects	X	X	X	X	X	X	X	X	X
Mean of Y	0.811	0.806	0.839	0.810	0.803	0.837	0.812	0.810	0.841

The probability of finishing compulsory school would equal 1 if the individual finished compulsory school or junior high school, 0 otherwise. As the age for finishing compulsory school is aging from 14 to 16, I also limited the cohorts from 1992 to 1987, where the affected cohorts were born from 1990 to 1992, and unaffected cohorts were from 1987 to 1989. The result is similar to the table above. All specifications include subdistrict and cohort fixed effects, and standard error is at the subdistrict level.
*** p<0.01, ** p<0.05, * p<0.1

Table 27:The Earthquake Effect on the Probability of Finishing High School

=1 if Finishing senior high school		All			Female			Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake X Exposed	-0.102**	-0.106**	-0.0601	-0.0433	-0.0597	-0.0582	-0.197*	-0.186*	-0.126
	(0.0462)	(0.0502)	(0.0476)	(0.0460)	(0.0516)	(0.106)	(0.103)	(0.106)	(0.0864)
Observations	2,901	2,697	1,700	1,518	1,411	848	1,383	1,286	852
R-squared	0.281	0.284	0.331	0.389	0.400	0.466	0.367	0.372	0.400
Personal Controls	-	X	X	-	X	X	-	X	X
Household Characteristics	-	-	X	-	-	X	-	-	X
Fixed Effects	X	X	X	X	X	X	X	X	X
Mean of Y	0.921	0.924	0.920	0.935	0.940	0.939	0.906	0.907	0.901

The probability of finishing compulsory school equals to 1 if the individual attended high school and graduated, 0 otherwise. As the age for finishing high school is aging from 17 to 19, I restrict the cohorts from the year 1984 to 1989, where the affected cohorts are born from 1987 to 1989, and unaffected cohorts are from 1986 to 1984. The result is similar to the table above. All specifications include subdistrict and cohort fixed effects, and standard error is at the subdistrict level.
*** p<0.01, ** p<0.05, * p<0.1

Table 28: The Earthquake Effect on the Probability of Working

Panel A: IFLS3 vs IFLS4	**=1 if Individual is Employed**			**=1 if Individual is Working in Construction**		
	All	Male	Female	All	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
EarthquakeXIFLS4	0.0780	0.337**	0.00466	0.0177*	0.0353*	0.00891
	(0.0902)	(0.156)	(0.0843)	(0.00901)	(0.0192)	(0.0110)
Observations	884	468	416	884	468	416
R-squared	0.277	0.419	0.380	0.202	0.248	0.504
Fixed Effects	X	X	X	X	X	X
Mean of Y	0.705	0.677	0.736	0.0124	0.0214	0.0024

Panel B: IFLS3 vs IFLS5	**=1 if Individual is Employed**			**=1 if Individual is Working in Construction**		
	All	Male	Female	All	Male	Female
	(7)	8)	(9)	(10)	(11)	(12)
EarthquakeXIFLS5	-0.276	0.00367	-0.616***	0.0189*	0.0312*	0.00771
	(0.254)	(0.411)	(0.223)	(0.00989)	(0.0177)	(0.00970)
Observations	909	483	426	909	483	426
R-squared	X	X	X	X	X	X
Fixed Effects	0.331	0.385	0.497	0.179	0.265	0.504
Mean of Y	0.596	0.592	0.601	0.0121	0.0207	0.00235

This specification includes age, community fixed effect, and survey wave fixed effect to count for unobservable characteristics due to location and time. Panel A compares the change between IFLS3 and IFLS4, showing the immediate effect of the earthquake. Panel B compares IFLS3 and IFLS5, showing the long-term effect of the earthquake. The sample consisted of individuals who were not in school or had worked before at the time of survey between age 15 and 18. The standard error is at the community level.

*** p<0.01, ** p<0.05, * p<0.1

Table 29: Nighttime Light Change and Economic Growth

	Annual Night-time Light Growth Rate (Change of Community Summmation)	Annual Night-time Light Growth Rate (Change of Community Median)
Pre-earthquake Period	-0.0215***	-0.127
2001 to 2004	(0.00648)	(0.124)
Earthquake Period	0.0129*	0.225*
2006	(0.00740)	(0.127)
Reconstruction Period	0.140***	2.148***
2007 to 2009	(0.0101)	(0.170)
After Reconstruction Period	-0.0231***	-0.444***
2010 to 2013	(0.00451)	(0.0761)
Observations	1,846	1,846
R-squared	0.623	0.508
Number of Communities	142	142
Fixed Effects	YES	YES
Mean of Y	0.0239	0.556

Regression result on the left measure the change in the total summation of nightlight points within 10km of the community radius and the regression results on the right measure the change in the community level night time light median. Robust standard errors are clustered at the community level. Fixed effects include year fixed effect and community fixed effect. The year fixed effect captures the unobservable characters that could affect the community nighttime change over time.

Table 30: Falsification Test

Years of Education	All			Female			Male		
	(1)	(2)	(3)	(7)	(8)	(9)	(10)	(11)	(12)
Earthquake X Exposed	-0.293	-0.349	-0.230	-0.390	-0.508	-0.569	-0.252	-0.276	0.212
	(0.274)	(0.281)	(0.643)	(0.483)	(0.435)	(1.034)	(0.495)	(0.514)	(0.867)
Observations	4,986	4,540	1,711	2,601	2,353	793	2,385	2,187	918
R-squared	0.420	0.441	0.526	0.503	0.525	0.623	0.431	0.458	0.556
Personal Control	-	X	X	-	X	X	-	X	X
Household Characteristics	-	-	X	-	-	X	-	-	X
Fixed Effects	X	X	X	X	X	X	X	X	X
Mean of Y	10.30	10.39	11.03	0.0489	0.0488	0.0559	0.0387	0.0376	0.0420

The specification follows the main analysis but defines the affected cohort differently. Instead of comparing school ages between 12 to 18, I increase the at-risk age by six years. Thus, my exposed group is aged 19 to 25, and the non-exposed group is aged 26 to 32. But they should be old enough, and their educational attainment not be affected by the earthquake.
*** p<0.01, ** p<0.05, * p<0.1

Table 31: Robustness Check Using Age 12 To 22 as Affected Cohort

Years of Education	All			Female			Male		
	(1)	(2)	(3)	(7)	(8)	(9)	(10)	(11)	(12)
Earthquake X Exposed	-0.630***	-0.614**	-0.535**	-0.495	-0.472	-0.0263	-0.822**	-0.826**	-0.964*
	(0.233)	(0.244)	(0.259)	(0.334)	(0.334)	(0.613)	(0.359)	(0.404)	(0.528)
Observations	5,649	5,097	2,702	3,029	2,720	1,341	2,620	2,377	1,361
R-squared	0.387	0.418	0.476	0.472	0.501	0.571	0.395	0.428	0.472
Personal Control	-	X	X	-	X	X	-	X	X
Household Characteristics	-	-	X	-	-	X	-	-	X
Fixed Effects	X	X	X	X	X	X	X	X	X
Mean of Y	10.69	10.78	11.09	10.64	10.74	11.11	10.74	10.83	11.07

The specification follows the main analysis but defines the affected cohort differently. Instead of comparing school ages between 12 to 18, I define affected cohorts as between ages 12 to 22, allowing time for college and other more advanced degrees.
*** p<0.01, ** p<0.05, * p<0.1

Table 32: Robustness Check Defining Earthquake as MMI Greater Than 5

Years of Education	All			Female			Male		
	(1)	(2)	(3)	(7)	(8)	(9)	(10)	(11)	(12)
Earthquake X Exposed	-0.697***	-0.629***	-0.213	-0.368	-0.274	0.384	-0.944**	-0.924**	-0.775*
	(0.208)	(0.225)	(0.278)	(0.381)	(0.405)	(0.766)	(0.412)	(0.465)	(0.455)
Observations	4,653	4,178	2,431	2,526	2,261	1,238	2,127	1,917	1,193
R-squared	0.399	0.431	0.487	0.495	0.526	0.597	0.400	0.434	0.470
Personal Control	-	X	X	-	X	X	-	X	X
Household Characteristics	-	-	X	-	-	X	-	-	X
Fixed Effects	X	X	X	X	X	X	X	X	X
Mean of Y	10.69	10.78	11.09	10.64	10.74	11.11	10.74	10.83	11.07

The specification follows the main analysis but defines an earthquake as MMI greater than 5.
*** p<0.01, ** p<0.05, * p<0.1

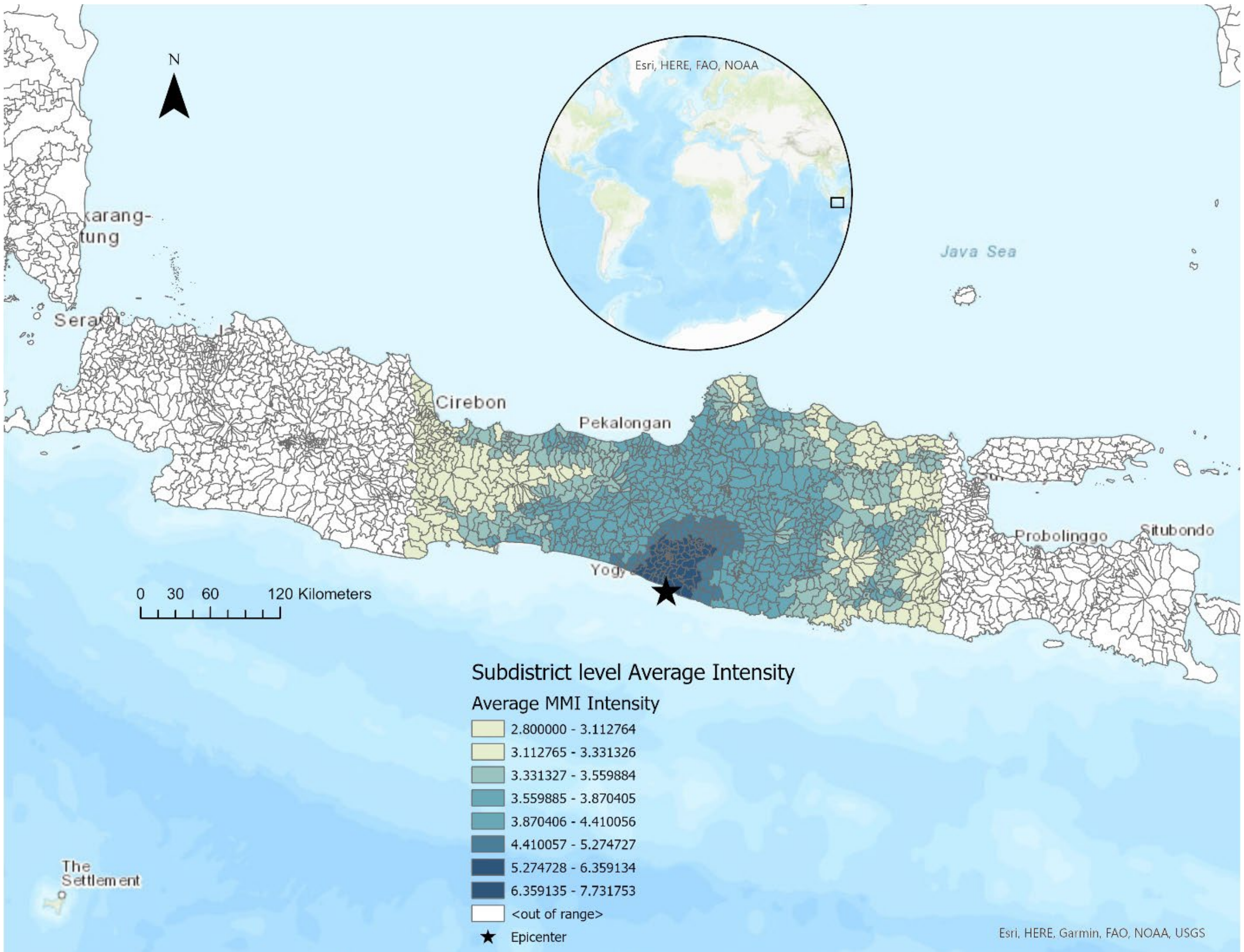

Figure 11: Modified Mercalli Intensity at the Subdistrict Level

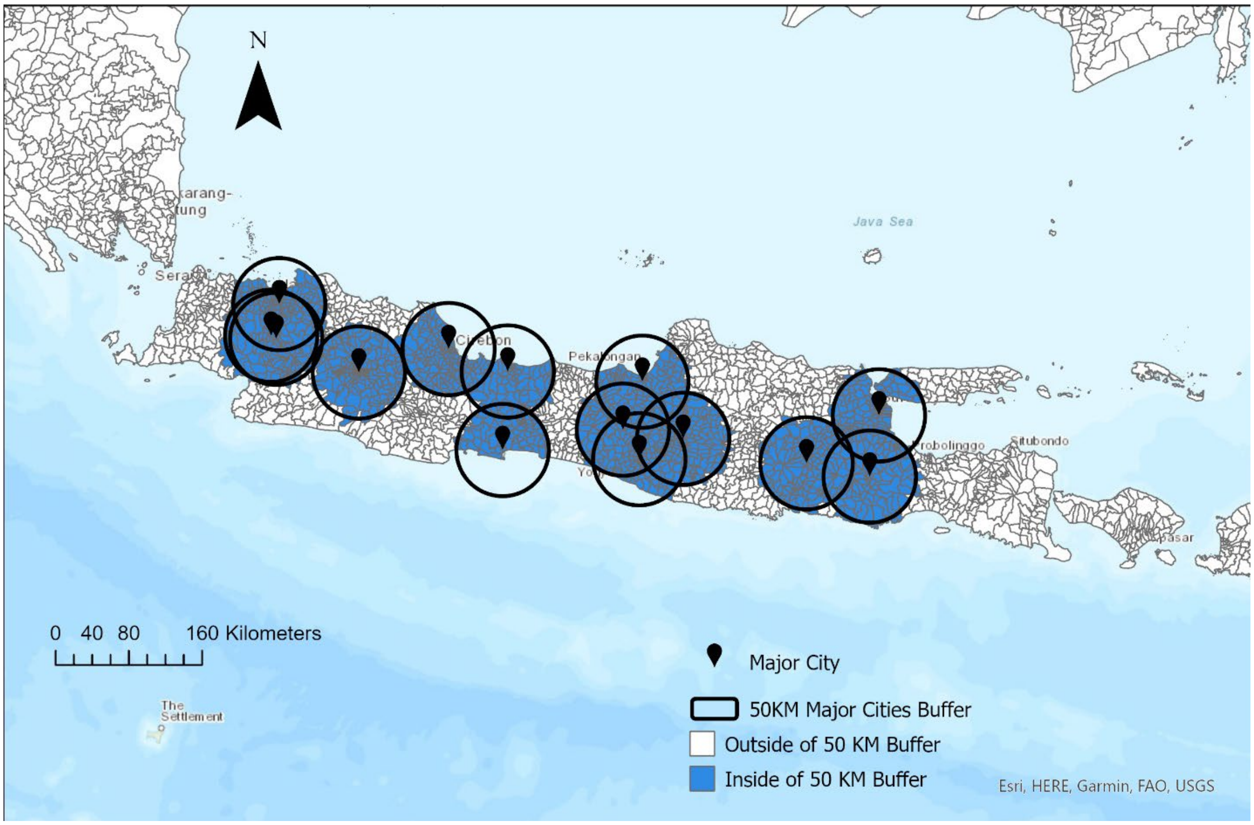

Figure 12: Major City Buffer and Subdistricts Selection

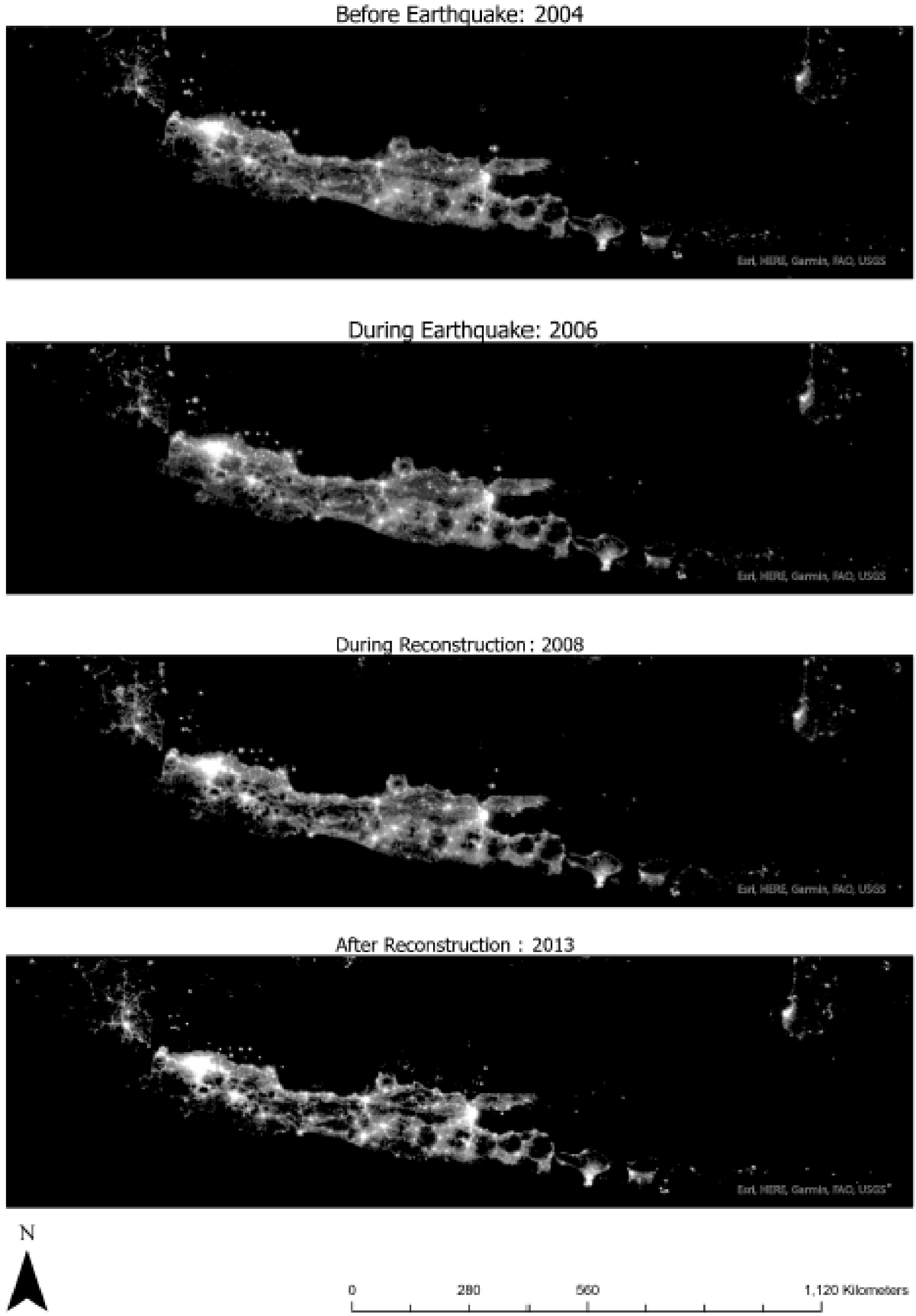

Figure 13: Nighttime Light in 2004, 2006, 2008, and 2013 (NOAA, 2021)

REFERENCES

Ahmed, K., Haq, S., & Bartiaux, F. (2019). The nexus between extreme weather events, sexual violence, and early marriage: a study of vulnerable populations in Bangladesh. *Population and Environment*, 40(3), 303-324.

Arya, S. and Agarwal, N., (2012). Prevention and control of infections after natural disasters. *Expert Review of Anti-infective Therapy*, 10(5), pp.529-530.

Alston, M., Whittenbury, K., Haynes, A. and Godden, N., (2014). Are climate challenges reinforcing child and forced marriage and dowry as adaptation strategies in the context of Bangladesh? *Women's Studies International Forum*, 47, 137-144.

Anderson, S. (2007). The Economics of Dowry and Bride price. *Journal of Economic Perspectives*, 21 (4): 151-174.

Asadullah, M. N., Islam, K. M., & Wahhaj, Z. (2020). Child marriage, climate vulnerability and natural disasters in coastal Bangladesh. Journal of Biosocial Science, 53(6), 948–967.

Ashraf, N., Bau, N., Nunn, N., & Voena, A. (2020). Bride Price and Female Education. *Journal Of Political Economy,* 128(2), 591-641.

Azage, M., Kumie, A., Worku, A., C. Bagtzoglou, A. and Anagnostou, E., (2017). Effect of climatic variability on childhood diarrhea and its high-risk periods in northwestern parts of Ethiopia. *PLOS ONE*, 12(10), p.e0186933.

Baird, S., McIntosh, C., & Ozler, B. (2011). Cash or Condition? Evidence from a Cash Transfer Experiment. *The Quarterly Journal Of Economics*, *126*(4), 1709-1753.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., & Sulaiman, M. (2020). Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa. *American Economic Journal: Applied Economics*, *12*(1), 210-259.

Barbut, F., Galperine, T., Vanhems, P., Le Monnier, A., Durand-Gasselin, B., Canis, F., Jeanbat, V., Duburcq, A., Alami, S., Bensoussan, C., & Fagnani, F. (2019). Quality of life and utility decrement associated with clostridium difficile infection in a French hospital setting. Health and Quality of Life Outcomes, 17(1).

Basu, A., & Gandhay, V. J. (2021). Quality-adjusted life-year losses averted with every COVID-19 infection prevented in the United States. Value in Health, 24(5), 632–640.

Bellizzi, S., Lorettu, L., Farina, G., Bubbico, L., Ferlito, S., Cegolon, A., Pichierri, G., & Cegolon, L. (2021). Humanitarian crises and child-marriage: Historical recurrent interrelated events. *Journal of Global Health*, *11*.

Blanco, L. M. B. (2023). Shifting Marriage Timing for Women: Destructive Events and Forced Displacement. *Working Paper*.

Bloom, D., Canning, D. and Sevilla, J., (2004). The Effect of Health on Economic Growth: A Production Function Approach. *World Development*, 32(1), pp.1-13.

Breierova, L., & Duflo, E. (2004). The Impact of Education on Fertility and Child Mortality: Do Fathers Really Matter Less Than Mothers?. *Working Paper*.

Brennan, R., (2005). Rapid health assessment in Aceh Jaya District, Indonesia, following the December 26 tsunami. *Emergency Medicine Australasia*, 17(4), pp.341-350.

Brown, R. and Velásquez, A., (2017). The effect of violent crime on the human capital accumulation of young adults. *Journal of Development Economics*, 127, pp.1-12.

Buchmann, N., Field, E., Glennerster, R., Nazneen, S., & Wang, X. (2021). A Signal to End Child Marriage: Theory and Experimental Evidence from Bangladesh. *Working Paper*.

Bush, K., O'Neill, M., Li, S., Mukherjee, B., Hu, H., Ghosh, S. and Balakrishnan, K., (2014). Associations between Extreme Precipitation and Gastrointestinal-Related Hospital Admissions in Chennai, India. *Environmental Health Perspectives*, 122(3), pp.249-254.

Caruso, G. and Miller, S., (2015). Long-term effects and intergenerational transmission of natural disasters: A case study on the 1970 Ancash Earthquake. *Journal of Development Economics*, 117, pp.134-150.

Caruso, G., (2017). The legacy of natural disasters: The intergenerational impact of 100 years of disasters in Latin America. *Journal of Development Economics*, 127, pp.209-233.

Cassar, A., Healy, A., & von Kessler, C. (2017). Trust, Risk, and Time Preferences After a Natural Disaster: Experimental Evidence from Thailand. *World Development*, *94*, 90-105.

Cdc.gov. (2020). *Infectious Disease*. [online] Available at: <https://www.cdc.gov/disasters/disease/infectious.html> [Accessed 23 November 2020].

Cdc.gov. (2020). *Keep Food and Water Safe After a Disaster*. [online] Available at: <https://www.cdc.gov/disasters/foodwater/facts.html> [Accessed 23 November 2020].

Cdc.gov. (2020). *Preparing for an Earthquake*. [online] Available at: <https://www.cdc.gov/disasters/earthquakes/prepared.html> [Accessed 23 November 2020].

Corno, L., Hildebrandt, N., & Voena, A. (2020). Age of Marriage, Weather Shocks, and the Direction of Marriage Payments. *Econometrica*, 88(3), 879-915.

Chari, A., Heath, R., Maertens, A., & Fatima, F. (2017). The causal effect of maternal age at marriage on child wellbeing: Evidence from India. *Journal Of Development Economics*, *127*, 42-55.

Chetty, R., & Looney, A. (2006). Consumption smoothing and the welfare consequences of social insurance in developing economies. *Journal of Public Economics*,90(12), 2351–2356.

Chicoine, L. (2020). Free Primary Education, Fertility, and Women's Access to the Labor Market: Evidence from Ethiopia. *The World Bank Economic Review*, *35*(2), 480-498.

Cutter, S., (2017). The forgotten casualties redux: Women, children, and disaster risk. *Global Environmental Change*, 42, 117-121.

Currie, J., Mueller-Smith, M. and Rossin-Slater, M., (2020). Violence While in Utero: The Impact of Assaults during Pregnancy on Birth Outcomes. *The Review of Economics and Statistics*, pp.1-46.

Duflo, E., Dupas, P., & Kremer, M. (2015). Education, HIV, and Early Fertility: Experimental Evidence from Kenya. *American Economic Review*, *105*(9), 2757-2797.

Dartanto, T. (2017). Disaster, Mitigation, and Household Welfare in Indonesia. Jakarta: LPEM-FEBUI.

Das, S., & Dasgupta, S. (2020). Marriage Market Responses in the Wake of a Natural Disaster in India. *Working Paper.*

Department for International Development. (2007). Development on The Record: DFID Annual Report 2007. *Annual Report 2007*.

Dong, Y. & Yang, H. (2020). Natural Disaster and Child Labor: Evidence from Indonesian Earthquake. *Working Paper.*

Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. American Economic Review, 91(4), 795-813.

El Arab, R., & Sagbakken, M. (2019). Child marriage of female Syrian refugees in Jordan and Lebanon: A literature review. *Global Health Action*, *12*(1), 1585709.

Eckstein, Z. and Zilcha, I., (1994). The effects of compulsory schooling on growth, income distribution and welfare. *Journal of Public Economics*, 54(3), pp.339-359.

Fabian, M., Lessmann, C. and Sofke, T., (2019). Natural disasters and regional development – the case of earthquakes. *Environment and Development Economics*, 24(5), pp.479-505.

Fan, X., Nie, G., Deng, Y., An, J., Zhou, J., Xia, C. and Pang, X., (2019). Estimating earthquake-damage areas using Landsat-8 OLI surface reflectance data. *International Journal of Disaster Risk Reduction*, 33, pp.275-283.

Field, E., & Ambrus, A. (2008). Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh. *Journal Of Political Economy*, *116*(5), 881-930.

Freeman, P., Martin, L., Mechler, R. and Warner, K., (2002). catastrophe and development. integrating natural catastrophe into development. *World Bank.*

Gadm.org. (2021). *GADM*. [online] Available at: <https://gadm.org/> [Accessed 25 February 2021].

Garcia-Hombrados, J. (2021). Child marriage and infant mortality: causal evidence from Ethiopia. *Working Paper*.

Gignoux, J. and Menéndez, M., (2016). Benefit in the wake of disaster: Long-run effects of earthquakes on welfare in rural Indonesia. *Journal of Development Economics*, 118, pp.26-44.

Gilleskie, D. and Hoffman, D., (2014). Health Capital and Human Capital as Explanations for Health-Related Wage Disparities. *Journal of Human Capital*, 8(3), pp.235-279.

Gupta, S. and Quick, R., (2006). Inadequate drinking water quality from tanker trucks following a tsunami disaster, Aceh, Indonesia, June 2005. *Disaster Prevention and Management: An International Journal*, 15(1).

Hallegatte, Stephane, Adrien Vogt-Schilb, Mook Bangalore, and Julie Rozenberg. Unbreakable: building the resilience of the poor in the face of natural disasters. World Bank Publications, 2016.

Heger, M. and Neumayer, E., (2019). The impact of the Indian Ocean tsunami on Aceh's long-term economic growth. *Journal of Development Economics*, 33, pp.275-283.

Heger, M. and Neumayer, E., (2019). The impact of the Indian Ocean tsunami on Aceh's long-term economic growth. *Journal of Development Economics*, 141, p.102365.

Ho, J., Chan, E., Lam, H., Yeung, M., Wong, C. and Yung, T., (2019). Is "Perceived Water Insecurity" Associated with Disaster Risk Perception, Preparedness Attitudes, and Coping Ability in Rural China? (A Health-EDRM Pilot Study). *International Journal of Environmental Research and Public Health*, 16(7), p.1254.

Jayaraman, A., Gebreselassie, T., & Chandrasekhar, S. (2008). Effect of Conflict on Age at Marriage and Age at First Birth in Rwanda. *Population Research And Policy Review*, *28*(5), 551-567.

Jafari, N., Shahsanai, A., Memarzadeh, M. and Loghmani, A., (2011). Prevention of Communicable Diseases After Disasters: A Review. *Journal of Research in Medical Sciences*, 16(7), pp.956-962.

James, B., (2020). *Disaster preparedness and mitigation: UNESCO's role*. [ebook] United Nation. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000150435> [Accessed 14 October 2020].

Kasozi, K., Namubiru, S., Kamugisha, R., Eze, E., Tayebwa, D., Ssempijja, F., Okpanachi, A., Kinyi, H., Atusiimirwe, J., Suubo, J., Fernandez, E., Nshakira, N. and Tamale, A., (2019). Safety of Drinking Water from Primary Water Sources and Implications for the General Public in Uganda. *Journal of Environmental and Public Health*, 2019, pp.1-12.

Keats, A. (2018). Women's schooling, fertility, and child health outcomes: Evidence from Uganda's free primary education program. *Journal Of Development Economics*, *135*, 142-159.

Kirchberger, M. (2017). Natural Disasters and Labor Markets. *Journal of Development Economics*, 125, 40-58.

Kling, J., Liebman, J. and Katz, L., (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, 75(1), pp.83-119.

Komarulzman, A., Smiths, J. and de Jong, E., (2014). Clean Water, Sanitation, and Diarrhea in Indonesia: Effect of Household and Community Factors. *Working Paper*.

Kouadio, I., Kamigai, T. and Hitoshi, O. (2011). Infectious Diseases Following Natural Disasters: Prevention and Control Measures. *Prehospital and Disaster Medicine*, 26(S1), pp.s125-s126.

Kumala Dewi, L. P., & Dartanto, T. (2018). Natural disasters and girls vulnerability: Is child marriage a coping strategy of economic shocks in Indonesia? *Vulnerable Children and Youth Studies*, *14*(1), 24–35.

Lemonick, D. (2011). Epidemics After Natural Disasters. *American Journal of Clinic Medicine*, 8(3).

Ligon, B., (2006). Infectious Diseases that Pose Specific Challenges After Natural Disasters: A Review. *Seminars in Pediatric Infectious Diseases*, 17(1), pp.36-45.

Mardani, M., (2018). Tetanus in Area Affected by Earthquake: Risk, Prevention and Management. *Archives of Clinical Infectious Diseases*, 13(5).

McGavock, T. (2021). Here waits the bride? The effect of Ethiopia's child marriage law. *Journal Of Development Economics*, *149*, 102580.

Nandi, A., Mazumdar, S., & Behrman, J. (2017). The effect of natural disaster on fertility, birth spacing, and child sex ratio: evidence from a major earthquake in India. *Journal Of Population Economics*, 31(1), 267-293.

Mosley, L., Sharp, D. and Singh, S., (2004). Effects of a Tropical Cyclone on the Drinking-Water Quality of a Remote Pacific Island. *Disasters*, 28(4), pp.405-417.

Neumayer, E., & Plumper, T. (2007). The gendered nature of natural disasters: The impact of catastrophic events on the gender gap in life expectancy, *1981-2002* . Annals of the Association of American Geographers. 97(3).

Nove, A., Matthews, Z., Neal, S., & Camacho, A. V. (2014). Maternal mortality in adolescents compared with women of other ages: Evidence from 144 countries. *The Lancet Global Health*, *2*(3).

Omar, Y., Parker, A., Smith, J. and Pollard, S., (2017). Risk management for drinking water safety in low and middle income countries - cultural influences on water safety plan (WSP) implementation in urban water utilities. *Science of The Total Environment*, 576, pp.895-906.

Paul, P. (2019). Effects of education and poverty on the prevalence of Girl child marriage in India: A district–level analysis. *Children and Youth Services Review*, *100*, 16–21.

Popkin, B., D'Anci, K. and Rosenberg, I., (2010). Water, hydration, and health. *Nutrition Reviews*, 68(8), pp.439-458.

Raj, A., Salazar, M., Jackson, E. C., Wyss, N., McClendon, K. A., Khanna, A., Belayneh, Y., & McDougal, L. (2019). Students and brides: A qualitative analysis of the relationship between girls' education and early marriage in Ethiopia and India. *BMC Public Health*, *19*(1).

Rathore, M., (2020). Infections after Natural Disasters. *Pediatrics in Review*, 41(10), pp.501-510.

Rautenberg, T. A., Downes, M., Kiet, P. H., Ashoush, N., Dennis, A. R., & Kim, K. (2022). Evaluating the cost utility of racecadotril in addition to oral rehydration solution versus oral rehydration solution alone for children with acute watery diarrhea in four low middle-income countries: Egypt, Morocco, Philippines and Vietnam. *Journal of Medical Economics*, *25*(1), 274–281.

Rheingans, R., Amaya, M., Anderson, J., Chakraborty, P., & Atem, J. (2014). Systematic review of the economic value of diarrheal vaccines. *Human Vaccines & Immunotherapeutics*, *10*(6), 1582–1594.

Romano, J. and Wolf, M., (2005). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica*, 73(4), pp.1237-1282.

Rundblad, G., Knapton, O. and Hunter, P., (2010). Communication, Perception, And Behaviour During A Natural Disaster Involving A 'Do Not Drink' And A Subsequent 'Boil Water' Notice: A Postal Questionnaire Study. *BMC Public Health*, 10(1).

Rumble, L., Peterman, A., Irdiana, N., Triyana, M., & Minnick, E. (2018). An empirical exploration of female child marriage determinants in Indonesia. *BMC Public Health*, *18*(1).

Rush, J. V. (2018). The impact of natural disasters on education in Indonesia. *Economics of Disasters and Climate Change*, *2*(2), 137–158.

Shemyakina, O. (2013). Patterns in Female Age at First Marriage and Tajik Armed Conflict. *European Journal of Population*, *29*(3), 303-343.

Strauss, J., F. Witoelar, and B. Sikoki.(2016). The Fifth Wave of the Indonesia Family Life Survey (IFLS5): Overview and Field Report. *WR-1143/1-NIA/NICHD.*

Sekhri, S., & Debnath, S. (2014). Intergenerational Consequences of Early Age Marriages of Girls: Effect on Children's Human Capital. *The Journal Of Development Studies*, *50*(12), 1670-1686.

Suprapto, Yanuarto, T., Nurmasari, R., & Prabowo, I..(2015). Population Exposed to Natural Hazard: a Study Based on the 2010 Population Census. United Nations Population Fund.

Tang, R., Wu, J., Ye, M. and Liu, W., (2019). Impact of Economic Development Levels and Disaster Types on the Short-Term Macroeconomic Consequences of Natural Hazard-Induced Disasters in China. *International Journal of Disaster Risk Science*, 10(3), pp.371-385.

Tsoukalas, D. and Tsitsifli, S., (2018). A Critical Evaluation of Water Safety Plans (WSPs) and HACCP Implementation in Water Utilities. *Proceedings*, 2(11), p.600.

UNICEF (2019). Saying No to Child Marriage in Indonesia. https://www.unicef.org/indonesia/stories/saying-no-child-marriage-indonesia.

UNICEF (2021). Child Marriage. https://www.unicef.org/protection/child-marriage.

United Nations Office for Disaster Risk Reduction (2009). Risk and Poverty in a Changing Climate: Invest Today for a Safer Tomorrow.

United Nation Office for Economic and Social Commission for Western Asia (2020). Challenges for Development in Current Conflict Settings: The Impact of Conflict on Child Marriage and Adolescent Fertility

United Nation, (2020). *UN Office for Disaster Risk Reduction Annual Report 2020 United Nations Office for Disaster Risk Reduction:2020 Annual Report*. UN Office for Disaster Risk Reduction.

USGS, (2021). *The Modified Mercalli Intensity Scale*. [online] Available at: <https://www.usgs.gov/natural-hazards/earthquake-hazards/science/modified-mercalli-intensity-scale?qt-science_center_objects=0#qt-science_center_objects>

Vásquez, W. and Aksan, A., (2015). Water, sanitation, and diarrhea incidence among children: evidence from Guatemala. *Water Policy*, 17(5), pp.932-945.

Vogl, T. (2012). Marriage institutions and sibling competition: Evidence from South Asia. *The Quarterly Journal of Economic.*128(3).

Vollaard, A. M., van Dissel, J. T., Surjadi, C., Visser, L. G., Widjaja, S., van Asten, H. A. G. H., & Ali, S. (2004). Risk factors for typhoid and paratyphoid fever in Jakarta, Indonesia. *JAMA*, *291*(21), 2607.

Vyas, S. and Kumaranayake, L., 2006. Constructing socioeconomic status indices: how to use principal components analysis. *Health Policy and Planning*, 21(6), pp.459-468.

Wado, Y.D., Sully, E.A. and Mumah, J.N. (2019) "Pregnancy and early motherhood among adolescents in five East African countries: A multi-level analysis of risk and protective factors," *BMC Pregnancy and Childbirth*, 19(1).

Wodon, Q., Nguyen, M. C., & Tsimpo, C. (2015). Child marriage, education, and agency in Uganda. *Feminist Economics*, *22*(1), 54–79.

Wang, J., Xu, W., Zhang, R. and Ni, W., (2009). Epidemic prevention and infection control of a field dressing station in Wenchuan earthquake areas. *International Journal of Infectious Diseases*, 13, p.S100.

Wald, D., Quitoriano, V., Heaton, T., Kanamori, H., Scrivner, C., & Worden, C. (1999). TriNet "ShakeMaps": Rapid Generation of Peak Ground Motion and Intensity Maps for Earthquakes in Southern California. *Earthquake Spectra*, 15(3), 537-555.

Waring, S. and Brown, B., (2005). The Threat of Communicable Diseases Following Natural Disasters: A Public Health Response. *Disaster Management & Response*, 3(2), pp.41-47.

Watson, J., Gayer, M. and Connolly, M., (2007). Epidemics after Natural Disasters. *Emerging Infectious Diseases*, 13(1), pp.1-5.

Woods, B., Revill, P., Sculpher, M., & Claxton, K. (2016). Country-level cost-effectiveness thresholds: Initial estimates and the need for further research. *Value in Health*, *19*(8), 929–*935*.

World Bank. (2023). *World Bank Open Data*. Retrieved February 25, 2023, from https://data.worldbank.org/

World Health Organization. (2003). *Making choices in health: WHO guide to cost-effectiveness analysis*. Tan-Torres Edejer T, Baltussen R, Adam T, Hutubessy R, Acharya A, Evans DB, Murray DB, Murray CJL, editors. Geneva: WHO.

World Health Organization. (2006). Communicable diseases following natural disasters: risk assessment and priority interventions. <https://www.who.int/diseasecontrol_emergencies/publications/who_cds_ntd_dce_2006.4/en/>

World Health Organization. (2011). Strengthening national health emergency and disaster management capacities and resilience of health systems. <https://apps.who.int/gb/ebwha/pdf_files/EB128/B128_R10-en.pdf?ua=1>

World Vision. (2013). Untying the knot: Exploring early marriage in fragile states.

Xu, K., Evans, D. B., Kawabata, K., Zeramdini, R., Klavus, J., & Murray, C. J. L. (2003). Household catastrophic health expenditure: A Multicountry analysis. *The Lancet*, *362*(9378), 111–117.

APPENDIX A: SUPPLEMENT MATERIAL FOR CHAPTER 1

In this section, I record the data cleaning process and the definition of the variables for this paper. This study has three layers of variables: personal level, household level, and communal level variables.

A.1 Indivudal Level Variables

At the individual level, my primary outcome variables are the incidence of acute symptoms. I define the disease variables equaling to 1 if the incidence of the diseases happened during the past four weeks by the time of the survey, and 0 otherwise. Although 15 different acute symptoms are reported in IFLS, I only record ten different acute symptoms as joint pain reported twice in the adult book, and the children's book does not contain certain variables. [38] The five waterborne disease symptoms are stomachache, nausea/vomiting, diarrhea, skin infection, and eye infection. And five none waterborne symptoms are headache, fever, toothache, cough, and running nose.

To control the personal difference that could have affected the incidences of the diseases, I added a vector of personal level control variables**. Age:** The age of the individual age at the time of the survey measured in years. **Age Square:** The square of the individual age at the time

[38] The 15 acute diseases are headache, runny nose, cough, difficult breathing, fever, stomachache, vomiting, diarrhea, painful or swollen joints, skin infection, eye infection, Toothache, painful or swollen joints, ear/nose/throat, kidney, heart/blood pressure, would/injury.

of the survey is used to measure the nonlinear relationship between age and incidence of the diseases. **Interview Month**: It records what month the interview happened. It controls the unobservable heterogeneity that changes over time within a year and affects the incidences of the diseases. **General Health**: It records four-level reported health conditions of each individual: 1 is very healthy, 2 is somewhat healthy, 3 is somewhat unhealthy, and 4 is unhealthy.

A.2 Household Level Variables

To explore why the reported health condition and incidence of waterborne diseases have changed, I create a vector of variables to track the accessibility to a safe environment factor in the household. First, IFLS asks the head of the household their **primary water source for drinking and using**. Ten possible answers are provided: aqua/air mineral, pipe water, well/pump (electric, hand), well water, spring water, rainwater, river/creek water, pond/fishpond, water collection basin, and others. Then, to define **safe water access**, I create an indicator that equals 1 if the household has access to pipe water or well/pump (electric, hand) and zero otherwise for drinking or using. Moreover, if a household has a water resource inside the house, the water inside of the house equals 1, and 0 otherwise. Three other variables are created to track the general living environment that could affect water safety. **Toilet:** It records if the household uses its toilets and this variable helps track if the household has access to toilet services rather than in the natural. It equals 1 if the household can access to any types of toilets. **Sewage:** Unsafe or unprocessed wastewater could result in water pollution and lead to waterborne diseases. **Garbage**: Like sewage, unsafely disposed waste could also result in water pollution and waterborne diseases. Thus, those two variables are used to control the household habits by the survey time, which equals 1 if the household uses sewage, flow or stagant, or access to garbaby service and 0 otherwise.

A.3 Community Level Variables

The earthquake's intensity is based on the Modified Mercalli Intensity (MMI) from the US Geological Survey (USGS), which estimates the overall impact of the earthquake. It gives a scale from I to XII, presenting an overall earthquake impact of the affected area. To link the earthquake intensity with the community, I use the supplement data from IFLS, which gives the exact GPS location of the community centers. By using ArcGIS-Pro, I can spatially link the communities with the earthquake intensity by GPS location.

A.4. Stepdown Multiple Hypothesis Test

When considering multiple hypothesis tests simultaneously, standard statistical techniques could over-rejection null hypotheses. Therefore, I calculate the stepdown adjusted p-values to correct the multiple hypothesis testing by following Romano and Wolf (2005). By setting a list of binary decisions concerning the individual null hypotheses as whole, this procedure helps to construct a better statistical estimation by using stepwise multiple testing procedure that asymptotically controls the familywise error rate. This procedure rejects more false hypotheses and captures the joint dependence structure of the test statistics, improving the ability to detect false hypotheses with different outcomes. Table B.1 and B.2 presents Romano-Wolf multiple hypothesis correction results. As this procedure considers the probability of rejecting at least one true null hypothesis in a family of hypotheses under the test, the overall results are more conservative than the original results. Although a less substantial impact is observed, this conservative procedure provides similar results as before.

Tables & Figures

Table A. 1: Romano-Wolf Step-Down Adjusted P-Values

Waterborne Diseases	IFL3 and IFLS4 Uncorrected P-value	IFL3 and IFLS4 Resample P-value	IFL3 and IFLS4 Roman-Wolf P-value	IFL3 and IFLS5 Uncorrected P-value	IFL3 and IFLS5 Resample P-value	IFL3 and IFLS5 Roman-Wolf P-value
Stomach	0.2694	0.2277	0.3564	0.2878	0.198	0.3663
Vomiting	0.4382	0.3564	0.3564	0.0082	0.0198	0.0198
Diarrhea	0.1609	0.0594	0.2277	0.4439	0.3069	0.3663
Skin Infection	0.0003	0.0099	0.0099	0.0000	0.0099	0.0099
Eye Infection	0.0027	0.0099	0.0099	0.0349	0.0297	0.0396

Non-waterborne Diseases	IFL3 and IFLS4 Uncorrected P-value	IFL3 and IFLS4 Resample P-value	IFL3 and IFLS4 Roman-Wolf P-value	IFL3 and IFLS5 Uncorrected P-value	IFL3 and IFLS5 Resample P-value	IFL3 and IFLS5 Roman-Wolf P-value
Headache	0.4750	0.2574	0.5446	0.0002	0.0099	0.0099
Running Nose	0.4264	0.1485	0.5446	0.2848	0.1485	0.3762
Cough	0.2058	0.0594	0.3663	0.3249	0.1782	0.3762
Diffcult Breathing	0.7654	0.6832	0.6832	0.5000	0.4455	0.4455
Fever	0.2089	0.1980	0.3663	0.0772	0.0495	0.1188

Living Environment	IFL3 and IFLS4 Uncorrected P-value	IFL3 and IFLS4 Resample P-value	IFL3 and IFLS4 Roman-Wolf P-value	IFL3 and IFLS5 Uncorrected P-value	IFL3 and IFLS5 Resample P-value	IFL3 and IFLS5 Roman-Wolf P-value
Safe Drinking Water	0.0005	0.0099	0.0099	0.0000	0.0099	0.0099
Water Inside House	0.184	0.1188	0.2475	0.4612	0.2574	0.7228
Safe Using Water	0.0297	0.0198	0.0396	0.0469	0.0099	0.0495
Have Toilet	0.8261	0.7723	0.7723	0.9982	1.0000	1.0000
Have Sewage	0.0006	0.0099	0.0099	0.6553	0.6238	0.8020
Have Garbage Serive	0.2963	0.0495	0.2673	0.2283	0.0099	0.3564

Note: The diseases incidence regression includes intensity, age, age square, the survey wave fixed effect, the interview month fixed effect, the individual fixed effect, and the standard errors are clustered at the community level. The number of resampling is 100 and seed is 50. The result is obtained through Stata command rwolf.

Table A. 2: Romano-Wolf Step-Down Adjusted P-Values

Waterborne Diseases	IFL2 and IFLS3 Uncorrected P-value	Resample P-value	Roman-Wolf P-value
Stomach	0.0002	0.0099	0.0099
Vomiting	0.0942	0.0693	0.0891
Diarrhea	0.1635	0.099	0.0999
Skin Infection	0.0003	0.0099	0.0099
Eye Infection	0.0393	0.0099	0.0396

Non-waterborne Diseases	IFL2 and IFLS3 Uncorrected P-value	Resample P-value	Roman-Wolf P-value
Headache	0.2356	0.099	0.2376
Running Nose	0.3068	0.1881	0.2376
Cough	0.038	0.0198	0.0594
Difficult Breathing	0.0166	0.0198	0.0297
Fever	0.0002	0.0099	0.0099

Living Environment	IFL2 and IFLS3 Uncorrected P-value	Resample P-value	Roman-Wolf P-value
Safe Drinking Water	0.738	0.5941	0.8515
Water Inside House	0.1459	0.1089	0.2871
Safe Using Water	0.362	0.2277	0.5347
Access to Toilets	0.4945	0.396	0.7327
Have Sewage	0.089	0.0297	0.1683
Garbage Service	0.7079	0.6832	0.8515

Note: The diseases incidence regression includes intensity, age, age square, the survey wave fixed effect, the interview month fixed effect, the individual fixed effect, and the standard errors are clustered at the community level. The number of resampling is 100 and seed is 50, and the result is obtained through Stata command rwolf.

APPENDIX B: SUPPLEMENT MATERIAL FOR CHAPTER 2

B.1. IFLS Data cleaning process

The key personal data is the age of the first marriage. We define child marriage as an individual entering any marital union before the age of 18. Therefore, we only look at the timing of the first marriage, and those marriages happened before reaching the age of 18. When available, we use the self-reported marriage age. However, some individuals reported the date of their marriage (year and month) but did not report their age at marriage. In these cases, we use the information on the date of marriage and date of birth to calculate the age at marriage and round down all ages to the lower bound. If only marriage year is reported, not marriage month, we use only the marriage year and birth year to calculate the age at marriage. If no information is available, we treat their age at first marriage as missing, if they are ever married. Thus, we drop those observations from our sample. In addition, as we study the effect of the earthquake on child marriage, we need to ensure respondents are at least 18 years old by the time of the last wave of the survey. Thus, we drop all individuals born after the year 1997 as they are not yet 18 by the time of finishing the last wave of the survey. As the key subgroup analysis used for this paper is their residetial location in the urban or rural area, we restrict we sample to individuals born after the year 1978 as they will be entering the "adult survey" in the first survey and have their residential locality reported by the household survey. In the case that same person reported different marriage information in different waves, we use the earliest wave information to diminishing the recall basis. Our final sample includes 20,250 unique individuals were found, of whom 1070 were already married but did not report their first marriage. We drop those who did not report their age at marriage. Also, we drop those who were married before 12 because we track the marriage time starting at the age of 12, and also because a reported age at marriage younger than 12 is likely to be an error. In total, 63 observations, counting 0.3% of total observations, were dropped because of this.

Another key information is residential location based on migration history. IFLS provides detailed individual migration information tracking the birthplace residential location, the residential location at 12, and any migration after 12. As we only study the marriage decision before 18, we recover the individual and their residential location history, in months, from the age of 12 to the last month of 17. By creating this panel information, we know where each individual lives for a given month after 12 and before 18. Sometimes, the migration month or the birthday month information is missing, and migration subdistrict information cannot be found. When the month information is missing, we treat all migration that happened in January. When the residential code is not reported, we use the birthplace subdistrict code to replace the missing subdistrict code if the birthplace subdistrict code is not missing. If the birthplace subdistrict is also missing, we treat the residential location as missing and drop those individuals from this study. In the case that same person reported different migration information in different waves, we use the earliest wave information to diminish the recall basis. Furthermore, between the first and fifth waves of IFLS, the administrative boundaries of some subdistricts changed. For instance, some subdistricts split into smaller ones, some subdistricts merged into bigger ones, and some subdistricts changed their political boundaries. As a result, the BPS codes used by Statistics Indonesia to identify the subdistricts and other administrative units in Indonesia may change over time. This might result in issues when matching earthquake with personal residential history. The detailed procedure to address this issue is discussed in the later section. As we do not have access to IFLS1 BPS code crosswalk, we do not use the migration history in IFLS1.

As the time of the first marriage do not have personal level variance and tracking household level variance for the time of marriages is difficult, we can only control the admirative level (subdistrict level) fixed effect. In this case, we create two sets of variables: personal control variables and before marriage household variables to control factors could have affected the time of the marriage in the case of earthquakes. For personal control variables, it includes age of each panel year, calculated by the birth year and the panel year[1], birth year, given by the survey data, religion, reported by the household head or individual survey, and ethnicity, reported by the household head or individual survey. For religion, it is reported by the household head at the time of the survey for each household individual across all waves. For ethnicity, it only started to be reported after the third waves (IFLS3 in the year of 2000), we use the individual reported ethnicity if it is available and use the household reported household main ethnicity if the individual level respondent is not available. We treat religion and ethnicity variables as time invariant.

We extracted seven household variables. The locality of the individual, which is reported by the household head at the time of the survey, either urban or rural area. As this locality could change over time, we followed the same procedure as Gignoux and Menéndez (2016) and assigned the first available data for that individual throughout the analysis no matter they moved or not at later surveys. Number of the children within the household is calculated by count of number of individuals younger than 18 living within the household at the time the survey.

[1]For example, if someone is born 1985 and the current panel year is 2001, the age is calculated as 16. The panel year is defined based on each individual's panel data from age of 12 and onward.

Number of unmarried siblings is counted how many of unmarried siblings are living within the household at the time of the survey. Birth order is calculated by rank the age of the children within the household at the time of the survey. Wealth quantile is based on the Principal Component Analysis (PCA) of three sets of variables: household infrastructure, durable household assets, and household living conditions. The household infrastructure includes indicators including accessibility to telephone, electricity, clean water supply, toilet, sewage, and garbage service. The household living conditions include the general living environment of type of house inside and outside wall, surrounding water, house stall, number of rooms, floor type, roof. The durable household assets include monetary value of the land owned, house, vehicles, jewels, furniture, and any other durable assets. We excluded all consumable assets such as food as they are very like to be time invariant. Then, we created a quintile of the first component of the PCA of those three sets of variables to indicate the socioeconomic conditions of the households where the first quantile indicates the poorest and the fifth quantile indicates the richest household at the time of each survey (McKenzie, 2005). At the time of any variable is missing, we take average of the year to replace the missing variable to ensure we have all household information. For mother and father's education, we extracted the information from the household level survey and created three indicators, less than primary education (less than 6 years of schooling or less than finishing elementary school), finished primary but not compulsory (less than 9 years of schooling or less than finishing junior high school), finished compulsory (finishing at least junior high school). After obtained all household variables, we filled the individual panel data based on the survey year. From 1993 to 1996 (IFL1 to IFLS2), we assigned the induvial in IFLS1 and unmarried with their household level data in IFLS1. From 1997 to 1999, we assigned the IFLS 2. Same procedure to IFLS3, IFLS4, IFLS5. Since we assume that household variable hold constant after the survey year until the next survey, we also assume that it held constant before the first year for another six years. In this case, we also assigned the IFLS1 household level information to year 1997 to 1993 for those who are unmarried at the time of the first survey. Thus, household variables only vary at time of the each survey.

B.2 Earthquake Data and Associated Earthquake Intensity

USGS provides detailed earthquake information including earthquake magnitude, focal depth, focal center to record the effect of the earthquake at very specific geographic level. Gignoux and Menéndez (2016) uses Centennial Earthquake Catalog data to estimate the Peak Ground Acceleration (PGA) and Modify Mercalli Intensity (MMI) based on Zhao et al (2006). Currently this dataset has been updated and replaced by another dataset provided by USGS at https://earthquake.usgs.gov/earthquakes/search/. We used ComCat to download all selected earthquakes using Python by inputting [-12, 12] Latitude and [80, 150] Longitude from 0:00:00 am January 1st, 1980, to 11:59:59 pm December 31st, 2014, to select all possible earthquakes that could have affected Indonesia during the study period. We use Anaconda to access the Combat database and use this follow code to download all associated earthquakes: getproduct shakemap shape.zip -s 1980-01-01 -e 2014-12-31-b 80 150 -12 -12 -d ~/tmp/shakemap-shapes. Note that since there are too many earthquakes, one line of code cannot download them at once. We split into different time frame to download all earthquake shapefiles. In total, we downloaded 1921 earthquakes and their associated earthquake shapefiles. The shape files include shapeMap data for all earthquakes. ShakeMap maps the earthquake intensity for each earthquake recoded by spatial polygons. As the files provide direct data on the earthquake intensity, no estimation on MMI distribution is needed as pervious literature (Gignoux and Menéndez, 2016). To define the effect and damage of the earthquakes, we followed Wald et al. (1999) and defined earthquake as

MMI higher than VI (or 6) since it would lead to damage on both the resistance and vulnerable buildings.

B.3 BPS Code Conversion Between Waves

IFLS BPS code changes due to 1) political boundary change, 2) large subdistricts split into smaller subdistricts, or 3) smaller subdistricts merged into new bigger subdistricts over time. In this case, when converting BPS code between each wave from IFLS2 to IFLS5 by using crosswalk provided by IFLS. However, the IFLS1 crosswalk is not provided by the survey as only kecid98 is available, not kecid93. We only convert the IFLS2 to IFLS5 in the process. Considering IFLS1 BPS code is nonconvertible, we decided not to use IFLS1 migration history in our analysis and assume that kecid98 indicates the BPS code in 1997 when the IFLS2 was conducted. In additional, in order to ensure we control for the subdistrict fixed effect, we created a new BPS code for all related subdistricts if they someone where related during the history survey. Table A.2 illustrates an example extracted from IFLS provided Crosswalk from IFLS2 to IFLS5. For example, we grouped the 3520060, 3520070, and 3520080 as new subdistrict as BPS 3891 since three of them were group into 3520071 in kecid14 (entry 4,5, and 6). 3520090 were never involved with any other subdistrict, which got assigned as a new subdistrict as BPS 3891. The last two entries are assigned as BPS 2314 as they were one group in kecid98 as 3520120. By having the new BPS code, we merged all related subdistricts as new subdistrict groups using the revised BPS code and created a new subdistrict unit for analysis.

B.4 GADM CC Code and IFLS BPS Code

To link the earthquake ShakeMap shapefile data and IFLS BPS code data, we needed another data source to connect those two data files. The Database of Global Administrative Areas (GADM) is database provides country administrative areas shapefiles and their associated regional code, cc_code. cc_code is the same as BPS code as they were used by the Badan Pusat Statistik to identify the geographic regions. By having the shapefiles and geographic regions codes together, we were able to link the earthquakes ShakeMap and IFLS migration location together to accurate identify the earthquake occurrence at specific location and time. However, as GADM was published in 2009 and IFLS4 uses the 2007 BPS code and IFLS5 uses the 2014 BPS code, the BPS code used by GADM could be different from the BPS code used by IFLS4 and IFLS5. Therefore, for consistency, we convert all GDAM code to IFLS5 BPS code if it is not already using IFLS5 BPS code. Due to some discrepancies, 21 subdistricts cc_codes did not have matched BPS codes from IFLS4 or IFLS5. Therefore, for those subdistricts, we manually check the names of the subdistricts from the crosswalk provided by IFLS and find the associated BPS codes for IFLS5 by a matched subdistrict name. The table A.3 includes information for such conversion. Note for Selat Sagawin or CC code 9108032, we cannot find any associated subdistricts in the IFLS5 crosswalk. Moreover, some special subdistricts such as reservoirs or forests had no one lived. Therefore, we were not able to find them in the IFLS. Those subdistricts were dropped in the later sections as no one lived in such subdistricts.

B.5 Link IFLS, GADM, and ShakeMaps together

The last stage of the data cleaning process to the link IFLS, GADM, and ShakeMap together. First, we created a new subdistrict shapefile by using the new assigned BPS code and merge GADM subdistricts to a new subdistrict shape file through cc-code and BPS code. All subdistricts were historically connected as illustrated in Figure 2. This new subdistrict shapefile provides a subdistrict shape file that is consistent in all IFLS waves. Then, we spatially connected the new subdistrict shapefiles with each of the earthquake ShakeMap shapefile by taking the arithmetic average of all connected MMI polygons within that subdistrict.

Specifically, ShakeMap shapefile contains polygons that record the intensity of the earthquake indicated by MMI. Then, when we spatially joined two shapefiles together, we created a shapefile indicates how many of the MMI shapefiles within that subdistrict boundary. By taking the athematic average of those overlapping MMI shapefiles, we calculate the average MMI overlapped with the subdistrict. For example, if an earthquake happened near a subdistrict and 5 MMI polygons overlaps within this subdistrict giving MMI as 5.5, 5.5, 5.6, 5.7 and 5.8. Then we assigned MMI 5.62 to the whole subdistrict. In other words, the data indicates that throughout the subdistrict, individuals lived in the subdistrict feel earthquake intensities varies from 5.5 to 5.8 depending how close they are to the earthquake. Assigning 5.62 MMI to the whole subdistrict is not an accurate measure considering the subdistrict is an area not a point. But this measurement provides a novel way to model the effect of earthquake in the economics literature to meet the needs of studying more than one earthquake or natural disasters events. After connected each earthquake with the new subdistrict shape file, we can accurately identify earthquake at each new assigned BPS level for a given time, at month level. Thus, our smallest time unit is a month, and the smallest geographic unit is a subdistrict. Multiple earthquakes could happen in the same subdistrict and same month. Therefore, if such multiple events happened within a month, we only keep the one with the largest MMI within that subdistrict month. Finally, we connected the earthquake data with each individual through their migration history using the new BPS code and time (year and month). However, as we only observe marriage at age yar, we aggregate all earthquake data to yearly level, keeping the largest earthquake MMI that year if multiple earthquakes happened.

B.6 Choosing different MMI to define earthquakes

This section includes table and results did not include in the main results section. Table B.1 repeat the same analysis for girls replicate the main results for girls choosing different MMI cutoff points. Note that, we define earthquake as any subdistrict with MMI higher than 6 in our study since MMI higher than 6 would lead to considerable damage to both vulnerable and resistant buildings. For our data cleaning process, we are very likely to underestimate the true effect of the earthquake intensity in our dataset as we take the average of overlap MMI polygons within the subdistrict boundaries and the higher earthquake intensity polygons are smaller. By taking average, the smaller intensity but larger polygons are more likely be overstated within the area, which underestimate true intensity or the effect of the earthquake. As such, we now change the definition of earthquake to 5 and 7. Table B.1 panel A shows the similar effect as the main results for girls, with smaller and imprecise effects. This is likely to be explained by the factor that lower-level intensity did not generate large enough effect to significant modify the household behavior but followed the same pattern as MMI VI. When we define earthquake as MMI higher than VII, the effect is different from main results. When earthquake intensity higher than VII, it indicates a havoc effect in the affected area. Such destructional effect is very likely to delay the marriage as major facilities and capitals were heavily damaged afterward.

B.7 Leading effect of the earthquake

To verify the validity of the empirical identification strategy, we employed a leading effect model to explore the effect of earthquake one year before it happened. The leading effect results indicate that the effect of earthquake decreased to 0.7 percentage points level, counting for 50% reduction in group mean and significant at 0.1 level, in the urban area. The effect diminished to insignificant level in the rural area despite a 57% increase in the annual hazard rate in the rural area. We argue that since the age, birth year, and marriage time are not accurately reported, measurement error in timing might explain those effects before the earthquake happened. The large reduction in the urban effect magnitude and insignificant effect in the rural area provides validity of the empirical model.

Tables & Figures

Table B. 1: MMI and Building Damage

		Building Damage	
MMI	Perceived Shaking	Resistant	Vulnerable
I-IV	Light	None	None
V	Moderate	Very Light	Light
VI	Strong	Light	Moderate
VII	Very Strong	Moderate	Moderate/Heavy
VIII	Severe	Moderate/Heavy	Heavy
IX	Violent	Heavy	Very Heavy
X-XII	Extreme	Very Heavy	Very Heavy

Note: The relationship between the intensity and MMI is estimated by Wald et al. (1999) using earthquakes that happened in California, US.

Table B. 2: Examples of IFLS BPS Crosswalk

kecid98	kecid00	kecid07	kecid14	BPS
3520060	3520060	3520060	3520060	3891
3520060	3520060	3520061	3520061	3891
3520070	3520070	3520070	3520070	3891
3520070	3520070	3520070	3520071	3891
3520080	3520080	3520080	3520071	3891
3520060	3520060	3520060	3520071	3891
3520080	3520080	3520080	3520080	3891
3520090	3520090	3520090	3520090	2311
3520100	3520100	3520100	3520100	2312
3520110	3520110	3520110	3520110	2313
3520120	3520120	3520120	3520120	2314
3520120	3520120	3520121	3520121	2314

Note: The Crosswalk is extracted from IFLS provided Crosswalk from IFLS2 to IFLS5. Kecid98 indicates the BPS code used for IFLS2, kecid00 is for IFLS3, kecid07is for IFLS4, and kecid14 is for IFLS5. The first two digits indicate province number, the third and fourth digits indicate district number, and the last three digits indicate the subdistrict number. The village BPS code is not reported for privacy purposes, even if they are interviewed during the survey.

Table B. 3: GADM BPS Code Matching Process with IFLS5

Province	District	Subdistrict	CC_3	IFLS5
Lampung	Lampung Barat	Bengkunat Belimbing	1801012	1813110
Lampung	Lampung Barat	Ngambur	1801013	1813090
Nusa Tenggara Timur	Belu	Botin Leobele	5306022	5321080
Nusa Tenggara Timur	Belu	Io Kufeu	5306023	5321050
Nusa Tenggara Timur	Belu	Kobalima Timur	5306041	5321120
Sulawesi Tengah	Morowali	Mori Utara	7203061	7212070
Sulawesi Tenggara	Konawe	Wawonii Tenggara	7403074	7412010
Sulawesi Tenggara	Konawe	Wawonii Timur Laut	7403083	7412030
Sulawesi Tenggara	Kolaka	Poli-Polia	7404042	7411030
Sulawesi Tenggara	Kolaka	Lalolae	7404051	7411080
Sulawesi Tenggara	Kolaka	Loea	7404052	7411030
Sulawesi Tenggara	Kolaka	Tinondo	7404082	7411100
Maluku	Maluku Barat Daya	Moa Lakor	8108040	8108042
Maluku Utara	Kepulauan Sula	Taliabu Barat Laut	8203062	8208080
Maluku Utara	Kepulauan Sula	Lede	8203063	8208070
Maluku Utara	Kepulauan Sula	Taliabu Selatan	8203064	8208020
Papua Barat	Raja Ampat	Selat Sagawin	9108032	dropped
Papua	Nabire	Siriwo	9404100	9410081
Papua	Pegunungan Bintang	Ok Aon	9417028	9417027
Papua	Pegunungan Bintang	Okbemta	9417063	9417056
Papua	Tolikara	Tagineri	9418122	9402225

Table B. 4: Defining Earthquakes using different MMI

Panel A: Define Earthquake as MMI higher than V

	All Individuals			**Urban**			**Rural**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake (MMI≥V)	-0.0001	0.000520	0.00940	-0.00342	-0.00330	-0.00307	0.0120	0.0122	0.0174
	(0.00486)	(0.00487)	(0.00719)	(0.00473)	(0.00483)	(0.00778)	(0.00907)	(0.00923)	(0.0107)
Observations	52,133	51,197	21,795	17,256	17,006	10,497	17,130	16,888	11,298
R-squared	0.030	0.030	0.029	0.019	0.020	0.026	0.038	0.039	0.039
N of subdistricts	1,520	1,515	631	602	598	395	515	508	378
Personal Controls	-	X	X	-	X	X	-	X	X
Household Controls	-	-	X	-	-	X	-	-	X
Mean of Y	0.0288	0.0281	0.0252	0.0128	0.0124	0.0137	0.0364	0.0365	0.0359

Panel B: Define Earthquake as MMI higher than VII

	All Individuals			**Urban**			**Rural**		
Earthquake (MMI≥VII)	-0.0110**	-0.0112**	-0.0119*	-0.00791	-0.00807	-0.0128*	-0.0157	-0.0166*	-0.0221*
	(0.00500)	(0.00502)	(0.00637)	(0.00566)	(0.00571)	(0.00684)	(0.00966)	(0.00989)	(0.0132)
Observations	52,133	51,197	21,795	17,256	17,006	10,497	17,130	16,888	11,298
R-squared	0.030	0.030	0.029	0.019	0.020	0.026	0.038	0.039	0.039
N of subdistricts	1,520	1,515	631	602	598	395	515	508	378
Personal Controls	-	X	X	-	X	X	-	X	X
Household Controls	-	-	X	-	-	X	-	-	X
Mean of Y	0.0288	0.0281	0.0252	0.0128	0.0124	0.0137	0.0364	0.0365	0.0359

Table B. 5: The Leading Effect of Earthquakes in the Previous Year

Panel A: Girls

	All Individuals			Urban			Rural		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake t= +1	-0.00262	-0.00294	0.00683	-0.00712*	-0.00730*	-0.00773*	0.0227	0.0225	0.0327
	(0.00671)	(0.00670)	(0.0101)	(0.00387)	(0.00413)	(0.00394)	(0.0244)	(0.0247)	(0.0312)
Observations	52,104	51,168	21,778	17,307	17,053	10,497	17,338	17,091	11,281
R-squared	0.030	0.030	0.029	0.019	0.020	0.026	0.037	0.038	0.039
Number of BPS	1,520	1,515	631	607	603	395	531	524	378
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal	-	X	X	-	X	X	-	X	X
Household	-	-	X	-	-	X	-	-	X
Mean of Y	0.0288	0.0281	0.0252	0.0136	0.0132	0.0137	0.0397	0.0398	0.0359

Panel B: Boys

	All Individuals			Urban			Rural		
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Earthquake t=+1	-0.000871	-0.00108	-0.000791	-0.000220	-0.000247	-0.000433	-0.00402	-0.00409	-0.00187
	(0.000689)	(0.000761)	(0.00108)	(0.000718)	(0.000775)	(0.000885)	(0.00435)	(0.00436)	(0.00294)
Observations	48,194	47,217	20,819	18,783	18,474	10,411	16,481	16,157	10,408
R-squared	0.005	0.005	0.005	0.002	0.002	0.005	0.008	0.009	0.010
Number of BPS	1,369	1,365	630	612	610	394	513	510	361
Fixed Effects	X	X	X	X	X	X	X	X	X
Personal	-	X	X	-	X	X	-	X	X
Household	-	-	X	-	-	X	-	-	X
Mean of Y	0.00344	0.00330	0.00240	0.00122	0.00114	0.000865	0.00388	0.00390	0.00394

Table B. 6

APPENDIX C: SUPPLEMENT MATERIAL FOR CHAPTER 3

Table C.1 records the major cities with more than 300,000 residents in Java based on the Atlas of Urban Expansion data in 2000.

Table C. 1: Major City GPS Coordinates

City Names	Latitude	Longitude
Bandung	-6.93	107.62
Bogor	-6.6	106.81
Cilacap	-7.69	109.03
Ciomas	-6.57	106.76
Cirebon	-6.7	108.5
Jakarta	-6.26	106.84
Kediri	-7.82	112.01
Magelang	-7.49	110.21
Malang	-7.95	112.63
Semarang	-7.02	110.4
Surabaja	-7.35	112.72
Surakarta	-7.58	110.8
Tegal	-6.92	109.08
Yogyakarta	-7.78	110.37