Measuring the Long-term Effects of Orphanhood

Nicholas Thomas Gardner

Professor Duncan Thomas & Professor Kent Kimbrough, Faculty Advisors

Honors thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University

Durham, North Carolina

2018

Acknowledgements

Special thanks to Professor Duncan Thomas, Professor Kent Kimbrough, and Professor Elizabeth Frankenberg for their guidance in the completion of this paper. I would like to further thank all members of the STAR lab of Duke University for their support and helpfulness throughout the past year.

Abstract

This paper works towards developing the narrative of orphans whose parent or parents died from natural disaster. By taking advantage of the unanticipated nature of death from the 2004 Indonesian tsunami, orphanhood can be treated as much closer to random than similar literature using data centered on HIV/AIDS related deaths. We use a community level fixed effects model to attempt to derive a causal relationship between orphanhood and both education and log wages. Our models suggest that orphaned males aged 14 and older at baseline complete 1-2 fewer years of education than their cohorts. The adverse effects persist in the long-term, as these orphans earn 26% less than non-orphan cohorts.

JEL Classification: I24; I25; I31; J24; J31

Keywords: Development Economics; Indonesia; Orphan; Education; Wages; Tsunami

I. Introduction

On December 26, 2004, the northern coast of Indonesia experienced one of the most economically devastating natural disasters in world history. A massive offshore earthquake caused a tsunami that resulted in 160,000 deaths and 4.5 billion USD worth of damages, which would amount to about 97% of Aceh's regional GDP. In the following years, the Indonesian government and supporting NGOs launched one of the most impressive reconstruction and recovery efforts to take place in a developing country. However, much of the damage done was irreversible. Approximately 10,000 young children (aged 9-18 at the time of the tsunami) lost one parent and are classified as "single orphans", while about 4,400 children lost both parents ("double orphans"); as a result, there was a sharp increase in the number of orphans in the tsunami-affected areas.

Using longitudinal panel data from the Study of Tsunami Aftermath and Recovery (STAR), we hope to gain some insight on long-term impacts of parental death. Orphanhood can have myriad effects on individual and household level outcomes. For example, prior research suggests that AIDS-related parental death can have adverse affects on school enrollment. Bicego et al (2003) find that orphans in Africa are less likely to be at their appropriate grade level than non-orphans, and Case et al. (2004) point towards within-household discrimination against orphans as a potential explanation for lower levels of education. Other research on non-AIDS related parental death has had varied results regarding orphans' schooling. Berg et al. (2014) do not find a statistically significant difference in failure rates between natural parental death and parental death due to external factors like accidents, but does find that orphans have higher failure rates and lower school performances than non-orphans.

Further literature suggests that parental death has measurable effects on outcomes such as height and early sexual debut. Beegle et. al (2006) find that maternal orphanhood is correlated with adolescent height deficiency, but the results are not conclusive as to whether these effects are permanent. Chae et al. (2013) note that orphan girls in Sub-Saharan Africa are more likely to have an early sexual debut and, in some cases, marry earlier than their non-orphan cohorts.

There are still many unanswered questions regarding parental death. With STAR data, we hope to address the question of whether parental death from natural disaster, rather than AIDS, yields a similar negative impact on completed levels of education. In most of the AIDS-related parental death research, it is often noted that orphans are systematically less wealthy than their non-orphan cohorts due to higher instances of AIDS in lower income communities. Previous STAR research suggests that socio-economic status is not a predictor of mortality within STAR data, and thus we do not expect orphans to be from poorer households than non-orphans (Frankenberg et al. 2011). Problems with parental death caused by disease will be discussed in more detail in the next section, but, unlike unexpected death from natural disaster, there exists strong evidence of adverse effects of AIDS on a household prior to a parent's death. As such, it is harder to draw out a causal relationship between orphanhood and any outcome of interest in these settings. However, STAR data provides information on parental death caused exclusively by an unexpected natural disaster, which may be much closer to "randomly assigned" than AIDs death. Thus, it may be possible to extract causal relationships. A more detailed explanation of what is available in the STAR data set is included below.

Frankenberg's research answers important questions about mortality from the tsunami, which serves as the framework for our discussion of parental death. She finds that there is a significant gender gap in mortality rates (more male survivors than female), which is most pronounced in communities closest to the coast. This likely implies that better swimmers or those who are physically stronger are more likely to survive. Contrary to prior research, height was a weak predictor of mortality, only significantly predicting survival for females. She finds no evidence of a link between education levels and mortality. The same holds true for quality of housing, overall suggesting socio-economic status is only a weak predictor of mortality. Specifically, socio-economic status is only significantly correlated with death for children and prime-age females, from which we draw the assumption that socio-economic status is not a predictor of parental death. However, Frankenberg does find evidence that parents often tried to save their children in the tsunami, as child death significantly increased parents' chance of death. This will have an important impact on whether we can call parental death "random" in this setting, and must be further explored to fully understand.

Furthermore, with extensive longitudinal data available, we can map out wage and consumption patterns to identify any possible relation between parental death and earning or spending patterns. This question has not yet been answered in contemporary research, primarily due to the lack of sufficient data. We find that older male orphans receive less education by as much as two years and earn relatively lower wages than their non-orphan cohorts, while these effects are not seen in female orphans, or younger orphans of both genders.

The following discussion and analysis will begin with an in-depth literature review. Many of the papers chosen will discuss AIDs-related parental death in order to provide a foundation for the shortcomings that an analysis of STAR data will work to overcome. Section III will break down how STAR data was collected and organized and provide some insight on the information and variables that are available. Section IV will build the theoretical framework of how we hope to claim that parental death from the tsunami can be treated as random. Finally, Section V will detail the empirical strategy of our regressions, provide summary statistics, and pivot into a discussion of all results.

II. Extended Literature Review

Anne Case et al (2004) provide one of the hallmark studies on parental death and its effects on education. Data was collected from 19 Demographic and Health Surveys (DHS) across 10 Sub-Saharan African countries between 1992 and 200. They find that orphans face a statistically significant drop in level of education even after controlling for household wealth. They offer many compelling approaches to the orphanhood problem that will be integral to the success of our research. First, the authors note that many orphaned children end up living with distant relatives or family friends, and are thus able to compare "within household" outcomes between the orphans and the biological children of their new caregivers. Case et al find strong evidence of within household discrimination against orphans. Their results also suggest that there are no gender effects, and that negative impacts on education are stronger among older children.

In addition, Case includes double orphanhood as an interaction term and not as the sum of individual parental death effects (as has been done in previous research). The authors find that this allows for more pronounced effects of double orphanhood. The results are not without their flaws, however, as all respondents were 14 or under, which may leave out effects unique to adolescents. Furthermore, the data cannot separate out the cause of parental death, which would be an interesting control variable in this analysis. While heavily drawing from Case's research, our STAR analysis will include a separate discussion of parental death among adolescents (14 and older). It may be the case that older children experience additional pressure to enter the workforce in order to support a household after the death of a parent. These unique effects can only be captured by distinguishing between adolescents and non-adolescents in the models.

One of STAR's fundamental strengths in analyzing parental death is the availability of longitudinal data. Thus, it is helpful to compare our goals with that of similar studies. Evans et al (2007) use data from rural Kenya collected in a primary school health program originally intended for treatment of intestinal infections. Using a 1998 questionnaire as a "pre parental-death" baseline, school participation data was collected at 4-5 randomly selected times in a year. Information on parental death was collected in a 2002 survey in 75 schools, using a proxy if the individual was not available. A subsample was crosschecked with an earlier mortality survey to gauge response accuracy. It should be noted that HIV/AIDS is the leading cause of parental death is Evans' data.

Evans et al. points out a fundamental weakness in longitudinal data collection, particularly in Sub-Saharan Africa: data availability and attrition. Of the over 11,000 eligible respondents, only 7,800 were included due to missing data from migration, etc. In contrast, STAR data boasts a low attrition rate of only 4%, with no existing evidence that this rate would be systematically higher among orphans. The authors also point out that, in using only cross-sectional data, results from Case et al (2004) likely suffer from omitted variable bias caused by lack of information on individual and household level characteristics of orphans. Similarly, Beegle (2005) did not include individual fixed effects or properly control for age and thus left room for unobserved heterogeneity.

When expressing the empirical effects of parental death, Evans et al. use a summation term to allow for possible time-compounding effects. In short, this captures and controls for the fact that not everyone's parent(s) died at the same time in this sample. Another potential problem from AIDS related death is the period of familial decay that occurs in the years prior to death. Treating an AIDS victim is costly and demanding, and children may have to reduce school participation to care for a sick parent, take care of things around the house, or earn extra income. (Bicego et al, Evans et al). To control for this, Evans uses indicator variables for up to 3 years before death. Other controls include age cohort-year gender indicators and pre-death household characteristics; their model includes an individual level fixed-effects variable.

Some of the issues addressed above will likely not carry into the STAR data set. For example, we will not include time-compounding effects as all parental deaths in the sample occurred on the same date. Furthermore, there fundamentally cannot exist a period of "decay" pre-death, as the tsunami was sudden and unexpected. We do, however, need to control for age and gender cohort indicators as well as pre-death household characteristics. Rather than using individual level fixed-effects, it makes more sense in the context of our model to use community fixed effects. Damage from the tsunami varies widely based on geography; most simply, the districts closest to the shoreline experience the most damage. We assume that individuals from high-damage communities are different from individuals from low-damage communities in unobserved ways (such as the stress from seeing more dead bodies), and so we use community level fixed effects. This fixed effect model will also allow for control over variation in funding received from post-tsunami aid programs. Additionally, many of the covariates that we are trying to measure benefit from community level fixed effects by controlling for the quality of a local school or school district, number of hospitals or clinics nearby, type and prevalence of local industry, etc.

In Evans' model, the parent death indicators are also interacted with individual, household, and community characteristics to test for other differential effects of parent death. The authors state, " for example, the magnitude of the parent death effect may depend on child age because older children are better labor market substitutes for parents, perhaps making them more likely to drop out of school after an adverse household income shock" (40). Also of concern are unobservable time-varying shocks such as local weather and crop prices. These specific examples are controlled for in their model, but there still may exist time-varying shocks that are not captured. As such, it will be important to carefully consider the possibility of unobservable variables or shocks in the STAR model.

Evans finds significant lower levels of school participation among orphans without evidence of recovery. Post parental death, participation is smaller by 5.5 percentage points with larger effects for maternal death than paternal. One possible explanation is that children are more likely to be fostered as a maternal orphan and fathers are more likely to be absent. There is no additional impact of becoming double orphan, but this may be due to small sample size. Further, younger children are more likely to drop out of school. Intuitively, older children who are still in school are positively selected on ability and those who show academic success are very unlikely to drop out. Finally, the authors find no gender differences, but inclusion of double interaction term suggests that young girls are more likely to decrease participation. No significant spillovers are found.

We may expect to find different results from the STAR data. As Evans notes, maternal orphanhood may have a stronger impact on school participation due to high rates of absentee fathers in Sub-Saharan Africa. If we find that this is not the case in Indonesia, we might expect the maternal and paternal effects on education to be more equal. However, Evans' intuition on why older children are less likely to drop out is crucial to our model and should still be considered.

Bozzoli (2016) offers yet another discussion of parental death based on longitudinal survey data. The data was collected in Cape Town, South Africa via the Cape Area Panel Study

(CAPS). The analysis focuses on parental death impacts on orphanhood. CAPS respondents were aged 14-22 at the time of the first interview. The author frames teen pregnancy as "a multifactorial phenomenon correlating with school drop-out, poverty, race and gender inequalities in decision-making". Furthermore, young female adults in AIDS-afflicted households are more likely to engage in transactional sex to supplement income.

The results support Bozzoli's speculations; female orphans younger than 17 (at the time of wave 1) are more likely to become a parent than non-orphans. These effects are muted in males. The author also models parental death against time and material investment from the primary caregiver. There is no compelling evidence that orphans experience less time investment from caregivers than nonorphans, but there is evidence that material investment is lower. These results are supported in other literature, as mentioned above. The author goes on to interact material investment with pregnancy and finds that the increase in odds of becoming a parent for orphans is larger for those with less material investment.

Bozzoli's research provides an important framework for any future analysis of fertility rates. STAR data is much more detailed than Bozzoli's and thus offers more room for creative and in-depth analysis. However, a proper discussion of marriage and fertility is outside the scope of this paper, which will instead focus on employment and wages as a proxy for well-being. Finally, our work offers the only existing attempt at isolating the effects of orphanhood from its cause (be it natural disaster or disease). Section IV will provide an in-depth explanation of how our estimates may be interpreted as causal.

III. Data

This analysis uses data from the Study of Tsunami Aftermath Recovery (STAR), led by Elizabeth Frankenberg of Duke University. The data is constructed from a collection of longitudinal surveys targeted at 30,000 individual respondents from 10,000 distinct households. Despite the large sample size, a low attrition rate of 4% was maintained throughout the 12-year surveying period. Among the age group of concern (6-18), 89% were interviewed by the final wave of surveys.

The first wave of surveys was conducted in May 2005, five months after the tsunami hit. The surveying period lasted approximately 12 months. Every member of the household was targeted; respondents age 11 or older would self-report if available, while younger children would have proxy responses. The sample was concentrated primarily in Aceh and North Sumatra. Districts, locally called "kabupaten", were selected based on proximity to a coastline that was affected by the tsunami; eleven kabupaten were located in Aceh and the remaining two were in North Sumatra. Note that "kecamatan", the unit of our community level fixed effects, is a subdivision of a kabupaten.

STAR data is comprised of multiple waves of surveys. STAR A was conducted in 2005, and four follow-up surveys were conducted annually. Two additional waves, STAR F and G, were held in 2014 and 2016 respectively. Questions from the RAND Corporation's 2004 SUSENAS provided the baseline for the questionnaires, but several questions specific to the tsunami were also included. The questionnaires in each wave were divided into books, with each intended for a particular respondent and targeted specific areas of concern. The following is a comprehensive breakdown of each book.

Book K: Titled "Control Book and Household Roster", Book K targets the original 2004 households. If respondents from the first household are found, they are asked basic questions about the well-being and current affairs of all existing members of the household.

Book 1: This book is addressed to the leading female household member and includes questions regarding household expenditures, financial transfers, pricing behaviors of necessities, and basic questions on quality of housing.

Book 2: Book 2 is addressed to the household head or the head's spouse and targets the household economy. More specifically, the book includes questions on agricultural and non-agricultural business, assets, borrowing behavior, and income.

Book 3: Book 3 was given to all household members aged 15 or older and contained a broad array of questions regarding education and employment, migration, physical and psychological health status, morbidities among a kinship network, marriage and fertility, and individual-level habitual behavior (i.e. food consumption or smoking). As discussed below, this particular analysis primarily uses data collected from Book 3 and Book 5.

Book 4: This book targeted all women aged 15-49 and focused entirely on fertility preference, including attitudes and usage of contraception, and pregnancy history.

Book 5: Respondents to Book 5 were aged 14 or younger, with children younger than 11 receiving proxy responses from a legal guardian. Similar to Book 3, questions focused on education, health, and morbidities.

Data for the analysis of orphans was obtained from Books 3, 4 and 5. In addition, data on older orphans was collected from Books 2 and 3 in order to capture wage and consumption. The analysis uses a dummy variable for orphanhood, which is given a value of 1 if respondents answered "yes" to the question "Did your mother/father die as a result of the tsunami?" In order to restrict the population to children aged 6-18, these observations are dropped if the age variable, collected from the covers of individual-level surveys, fell outside the given range. Data on the respondent's location in wave A is gathered in order to control for community specific fixed effects, and individuals are dropped from the data if they are from a district that experienced no mortality from the tsunami. Education information was collected using book 3 and 5 responses that report years of primary, secondary, and tertiary education. These variables will be summed to create a single "years of education" variable. By design of the survey, each respondent receives a unique ID number, which was used to merge all the observations above onto one data set.

The strength of the variable here may come into question. Particularly, there may be a concern that self-reported results decline in accuracy as education level declines. However, we assume that education does not affect an individual's awareness of parental mortality. The clarity of the survey should allow us to assume there is no artificial correlation between orphanhood status and education level due to measurement error.

Information on household characteristics in wave A is collected using responses from questions on household size and expenditure. Additional controls include age, household size and years of education of mothers in wave A, which are available in the survey books listed above. It is clear that the STAR data set is comprehensive enough to carry out the multi-dimensional analysis of parental death that I have laid out thus far. The level of specificity available in these surveys allows for a much more nuanced discussion of economic outcomes for orphans other than education, which has been the centerpiece of much of the literature to date.

IV. Issues of Causality and Unobserved Heterogeneity

Before moving on to empirical analysis, it will be important to establish the theory and assumptions on which the upcoming discussion will be built. Much of the literature discussed thus far centers on parental death as a result of HIV/AIDS or other diseases. It has further been established that disease related orphanhood is subject to causality concerns; for example, the period of familial decay mentioned by Bicego will have a real and significant impact on outcomes like education. These factors of unobserved heterogeneity are difficult, if not impossible, to control for in the regressions discussed in the above literature. As such, the models cannot fully imply causality.

However, the question remains if we can find a causal relationship from parental death as a result of natural disaster. It is clear that we will not encounter problems with pre-disaster familial decay, as the tsunami came without warning. Thus the only concern when establishing a causal relationship is whether tsunami deaths were randomly "assigned". What, if any, are the significant determinants of tsunami mortality? If we find significant predictors of mortality, we must control for pre-tsunami household characteristics in order to wipe our model of unobserved heterogeneity.

The following table is taken from Ava Cas' work on parental death from earlier waves of STAR data. Using data from the first wave of STAR surveys, Cas compares certain indicators to determine if parental death is exogenous.

Table 1: Child and family characteristics at pre-tsunami baseline stratified by survival status of parents

	Both Parents Survived Tsunami	One or Both Parents Died During Tsunami	Difference (2) - (1)	Mother Died	Father Died	Both Parents Died
	(1)	(2)	(3)	(4)	(5)	(6)
Child Characteristics (pre-tsunami)						
Age (years)	12.9	13.5	0.6	0.6	0.2	0.8
	(0.1)	(0.2)	(0.3)	(0.4)	(0.5)	(0.4)
Male (%)	54.2	65.5	11.3	8.2	8.2	19.9
	(2.1)	(4.7)	(5.1)	(8.3)	(9.1)	(8.8)
Education (years)	5.3	6.4	1.1	1.1	0.9	1.5
	(0.2)	(0.2)	(0.3)	(0.4)	(0.6)	(0.5)
Enrolled in School (%)	91.4	96.6	5.3	5.0	5.5	5.4
	(1.5)	(1.5)	(2.2)	(2.8)	(3.5)	(3.6)
Working for a Wage (%)	4.6	2.8	–1.8	–2.7	3.1	–4.6
	(1.6)	(1.5)	(2.2)	(2.5)	(5.5)	(1.6)
Engaged in Housekeeping (%)	8.2	3.8	–4.4	–4.3	–8.2	–1.3
	(1.8)	(2.9)	(3.4)	(4.2)	(1.8)	(5.1)
Family Characteristics (pre-tsunami)						
Mother's Education (years)	8.5	8.6	0.1	0.5	–0.8	0.3
	(0.4)	(0.4)	(0.5)	(0.8)	(0.8)	(0.6)
Father's Education (years)	9.4	9.0	–0.4	–0.5	–0.9	0.3
	(0.4)	(0.4)	(0.5)	(0.7)	(0.8)	(0.7)
Mother Alive at Baseline (%)	98.1	99.2	1.0	1.9	–1.3	1.9
	(0.7)	(0.9)	(1.0)	(0.7)	(3.2)	(0.7)
Father Alive at Baseline (%)	95.4	90.8	–4.7	–0.8	–4.8	–11.6
	(1.0)	(4.0)	(4.0)	(5.3)	(9.0)	(7.3)
Per Capita Expenditure	40.0	41.8	1.8	7.9	–10.1	3.1
(Rp 10,000 per month)	(2.7)	(5.7)	(5.9)	(11.0)	(3.7)	(5.6)
Household Size	5.9	5.8	–0.1	0.0	–0.1	–0.5
	(0.2)	(0.2)	(0.2)	(0.3)	(0.4)	(0.3)

* (adapted from Cas et al.)

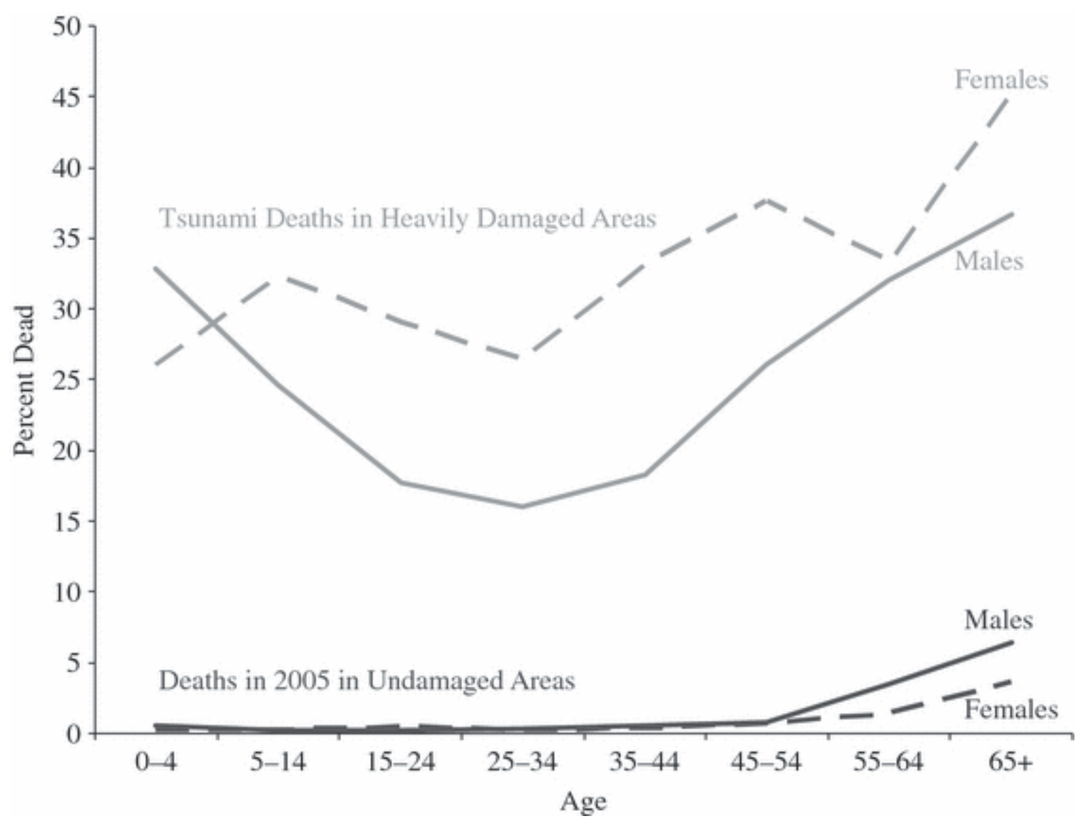
** Robust standard error in parentheses, adjusted for clustering

*** 590 children were interviewed whose parents both survived (1), and 119 children were interviewed who lost one or both parents (2)

The first three rows of the table report age in years, the percentage of respondents who are male, and years of completed education. Columns 3-6 report the difference for each of these variables between non-orphans and orphans (categorized by single orphan, maternal orphan, paternal orphan, and double orphan, respectively). All of these differences, for all three variables, are positive. In summary, Cas et al. find that children who survived the tsunami and lost at least one parent are older, more likely to be male, and are more likely to be enrolled in school with higher levels of schooling than survivors who did not lose a parent. The fact that orphans were positively selected on education may allow us to consider orphanhood as the driving factor in our models if we see a fall in education level among orphan. We expect some of this difference in level of schooling to be due to the age gap between the two groups. However, upon inclusion of a kecamatan fixed effect (community-level fixed effect), these differences are no longer statistically significant. We infer, then, that orphans are not statistically different than their cohorts within a community before the tsunami. By controlling for pre-tsunami household characteristics, we hope that these results reported by Cas et al. allow us treat our results as unbiased and give a causal relationship to parental death on education (Cas, 2014).

Some of Cas' findings on mortality are supported in other STAR research. Frankenberg et al. (2011) led research on mortality rates compared across gender, physical strength, socio-economic status, and household composition. The authors find that women were more likely to die in the tsunami than men (see Figure 1). Among women, those who were shorter were more likely to die in the tsunami than tall women. This height relationship is not present among men. Furthermore, they find no correlation between mortality and levels of education or occupation. The only indicator of socio-economic status that was significant was household ownership, which had a positive relationship with mortality for children and prime-aged females. However, demographic composition of the household seems to have a tangible effect on mortality. For example, the presence of a prime-aged male within a household increases the probability of survival, particularly among children and women. This result challenges the assumption that the gender mortality gap is due entirely to differences in physical strength, as many men helped save the weaker members of their household (Frankenberg, 2011).

Figure 1.



(Source: Frankenberg et al.)

The results pushed forward by Frankenberg et al. present a complicated story of mortality as a result of the tsunami. Knowing that education and occupation are predictors of mortality, we may assume that, when controlling for these variables and other pre-tsunami household characteristics and comparing only those children from kecamatan that experienced some levels of mortality, parental death can be treated as randomly assigned. We control for pre-tsunami household size to work towards controlling some of the variance caused by certain household members attempting to save others. These precautions should allow us to interpret our results as causal.

V. Results

The empirical strategy and corresponding results in this section will be presented through the two major segments of the analysis: education and wages.

A. Education

Our first and most basic model will focus on education, and should map years of completed education against the variable of concern, called "Orphan" here, which represents an indicator taking value 1 if at least one parent died in the tsunami. Dummy variables for each age at the time of the tsunami (7-18) are added to flexibly control for variation across time. The model is estimated allowing for clustered residuals and is represented as follows, with the subscripts for *A* representing the specific value for age:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 Orphan_i + \omega Z_i + \gamma A_7 + \cdots + \delta A_{18} + \mu_i} \qquad \textbf{(1)}$$

In this model, the constant β_0 represents the average completed level of education for children aged 6 that did not lose a parent in the tsunami. The covariate Z_i represents a vector of control variables for individual *i*; for these analyses, pre-tsunami household size and mother's education will be used as controls. These controls were selected based on the prior STAR literature on mortality, as discussed above. Additionally, we hold the assumption that unobserved heterogeneity μ is not correlated with *Orphan*.

It is crucial to flesh out exactly what population is represented in our model. More specifically, in order to control for any fixed effects rooted in exposure to the tsunami, we should only compare orphans to children who lived in a kecamaten with at least one death as a result of the tsunami. The fixed effect term will also sweep out any variance caused by community specific, post-tsunami recovery efforts. This fixed effect ε is separated out from the error term μ in model 2 below.

In order to allow for a more flexible interpretation of our results, we also separate out our parental death dummy into an indictor for the death of one parent and another for the death of both parents. Although we have controlled for age as a covariate, we will estimate two additional models that stratify the sample by reporting results for children younger than 14 and a second model for those 14 or older. Additionally, we will include a gender dummy and interaction term so as to capture variation across gender and to check for any statistically significant difference between genders. This new model is as follows, with covariates estimated using robust standard errors.

$$\boldsymbol{Y_i = \beta_0 + \beta_1 Orphan_{single} + \beta_2 Orphan_{double} + \omega Z_i + \gamma A_7 + \cdots + \delta A_{18} + \varepsilon + \mu_i} \qquad \textbf{(2)}$$

Table 2 below includes the coefficient estimates for the orphanhood variable in 8 separate models. The models are split up as follows: four OLS estimates that consider all orphans, only males, only females, and a difference estimate, and four community fixed-effect models (Kec FE) that follow the same gender control pattern. Results for both model (1) and (2) above are reported here; the first half of the table specifies the effect of at least one parental death, while the latter half controls for both single and double orphans. The results are presented with t-statistics in brackets and important results highlighted in red. It is important to note that each coefficient in the table is from its own regression; in other words, separate regressions were run for all ages + all genders, over 14 and male only, etc. Furthermore, each coefficient represents the equivalent of $\boldsymbol{\beta_1}$ or $\boldsymbol{\beta_2}$ described above. The control variables are not listed here, as coefficients for household size and mother's education were generally not significant, as expected.

Table 2: Comparing Completed Years of Education between Orphans and Same-aged Cohorts

[t stat]	OLS All	OLS Males	OLS Females	OLS M-F Diff	Kec FE All	Kec FE Males	Kec FE Females	Kec FE M-F Diff
Completed years of education								
Orphan (mother, father or both died)								
Age 6-18 yrs at baseline	-0.65*	-1.28***	0.41	-1.69**	-0.34	-1.02***	0.74	-1.76***
	[1.66]	[2.97]	[0.61]	[2.20]	[0.92]	[2.48]	[1.12]	[2.42]
Age 6-13 at baseline	-0.49	-0.66	-0.10	-0.56	-0.47	-0.74	0.06	-0.81
	[1.06]	[1.14]	[0.12]	[0.53]	[0.90]	[1.36]	[0.07]	[0.90]
Age 14-18 at baseline	-0.81	-1.97***	0.94	-2.92***	-0.28	-1.47***	1.37	-2.84***
	[1.44]	[3.28]	[0.95]	[2.52]	[0.60]	[2.84]	[1.47]	[2.59]
One parent or both parents died								
Age 6-18y at baseline								
One parent died	-0.36	-1.25***	0.93	-2.18	-0.09	-0.99	1.21***	-2.20***
	[0.83]	[2.18]	[1.46]	[2.53]	[0.23]	[1.76]	[2.34]	[2.88]
Both died	-1.38***	-1.35***	-1.61	0.25	-1.07*	-1.11**	-1.18	0.07
	[2.13]	[2.17]	[1.08]	[0.17]	[1.68]	[2.35]	[0.77]	[0.05]
Age 6-13 at baseline								
One parent died	-0.28	-0.82	0.59	-1.41	-0.41	-1.07	0.64	-1.70*
	[0.61]	[1.13]	[0.73]	[1.15]	[0.89]	[1.58]	[0.93]	[1.71]
Both died	-1.21	-0.25	-4.46*	4.21*	-0.70	0.08	-3.69	3.77
	[1.00]	[0.27]	[1.79]	[1.86]	[0.43]	[0.07]	[1.20]	[1.54]
Age 14-18 at baseline								
One parent died	-0.46	-1.78***	1.35	-3.13***	0.10	-1.05	1.61*	-2.66***
	[0.68]	[2.28]	[1.38]	[2.48]	[0.18]	[1.64]	[1.69]	[2.31]
Both died	-1.51*	-2.31***	-0.11	-2.20	-1.10*	-2.25***	0.73	-2.98*
	[1.85]	[2.58]	[0.07]	[1.20]	[1.67]	[2.75]	[0.53]	[1.79]

t-statistic in parentheses
* p < .1, ** p < .05, *** p < .001

By design of the model, the coefficient estimates above should be interpreted as the difference in completed years of education between an orphan and a same-aged, non-orphaned cohort. For example, orphaned males aged 14-18 are about 2 years behind their non-orphaned cohorts. These results form the foundation of the orphanhood problem.

Our analysis of these results will focus on differences that arise through gender, age, and single vs. double orphanhood status. The models show relatively consistent, significant differences between years of education for males and females, with males typically completing fewer years of education. The exception to this trend is the young female OLS group, who completed 4.5 fewer years of education than their cohorts. This relationship is muted in our fixed effect model, suggesting that this can be attributed to variation across communities or a result of a small sample size, which will become a consistent roadblock in our discussion. However, the male-female difference tends to be largest and most significant for the age 14 and older bucket, particularly when controlling for community fixed effects. In the third row and final column of Table 2 we capture the male-female difference in this age bucket without controlling for single or double orphanhood. Note that the difference is significant, but the coefficient for females is not. This suggests that all the variance in the male-female difference in education among orphans is derived from the males, who pursue about 1.5 fewer years of education than non-orphans (without controlling for single vs. double orphanhood).

The models show no statistically significant results for orphans younger than 14, with the exception of the female OLS group listed above. Among older children, we find slightly different effects. When one parent died, males pursued one fewer year of education on average when controlling for fixed effects. This effect is more than doubled when both parents died, suggesting that the burden of responsibility to replace a household head (and thus drop out of school) for elder male orphans is significantly larger among double orphans. Additionally, we see that older girls complete about 1.5 extra years of education than their non-orphaned cohorts, but this is only true in the single orphan case, which may be caused by a small sample size, or could be indicative of some sort of omitted variable bias. Female double orphans do not show any change in completed years of education.

Some aspects of our results are corroborated by prior literature, while others offer new dimensions to the orphan condition. Case et al. (2004) noted that there are no gender effects for education outcomes among orphans, and that negative impacts on education are stronger among older children. Our results suggest that, in the context of parental death from natural disaster, there may exist significant gender differences while supporting the idea that older orphans are more adversely affected. Other prior literature such as Bicego et al. (2003) and Berg et al. (2014) similarly suggest that orphanhood adversely affects educational attainment.

With the observed large and significant, negative effects on older male orphans in mind, we will need to properly control for education in our employment models so as to isolate the pure orphanhood effect. However, it appears that when a parent dies, the burden of responsibility typically falls on the oldest male child to fill the void. This effect is much more pronounced when both parents die; it is likely the case that these male children are expected to drop out of school in order to work and provide income to the household.

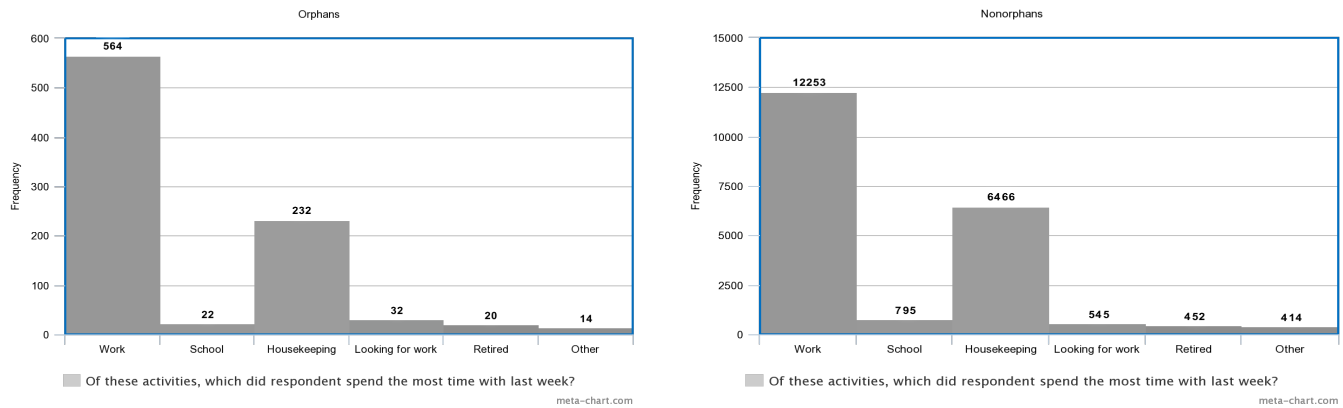
B. Wages

Moving forward from education, we will expand our discussion to other markers of socio-economic status. Our model for employment outcomes will require a slightly different specification than that used above. Data from the STAR survey was used to construct a set of indicator variables, *Working*, *School*, and *Neither*, which indicate whether an individual is primarily engaged in the labor market, school, or neither. It is important to note that “primarily engaged” here means that the respondents spends the most time doing this activity relative to other options (including taking care of the household, etc). Using this information, we will pivot into a model that captures the effect of orphanhood on log wages that is conditional on our indicator *Working* equal to one. The model will follow the same structure as model (1) above, however the orphanhood variable will be split into male and female orphan indicators to capture any gender differences. As before, separate models will be run for ages 7-13 and 14-18 to capture any adolescent-specific effects. Note that the control vector $\boldsymbol{Z}$ now includes years of education.

$$\boldsymbol{ln(wage) = \beta_1 Orphan_i + \omega Z_i + \gamma A_7 + \cdots + \delta A_{18} + \varepsilon + \mu} \qquad \textbf{(3)}$$

The following figure details the data available on respondents' employment. Within the employment STAR survey, respondents were asked what they spent the most time doing over the course of the previous week. The strength of using last week's activity as a proxy for yearly employment information may come into question. However, the survey asks respondents who did not report working during the previous week if they are actually employed. A vast majority of respondents answered no, and there is no significant difference between orphans and non-orphans in regards to the strength of this proxy. The following table compares distributions of responses between orphans and nonorphans. While the distributions are similar, 64% of orphans reported working compared to the 58% of non-orphans. Similarly, 26% of orphans reported primarily engaging in housekeeping compared to 31% of nonorphans.

Figure 2



Similar distributions were run for other employment variables and histograms are included in Appendix A. The results are summarized as follows: 44% of orphans report as self-employed, which is greater than the 38% of non-orphans. This difference is primarily due to a greater percentage of non-orphans self-reporting as unpaid family workers. The percentages of respondents who are government workers or privately employed are nearly identical. Additionally, when we compare across industry we find that orphans are less likely to be engaged in agricultural work than non-orphans by about four percentage points. When controlling for whether a respondent is currently working, it may be of interest to compare wages between orphans and non-orphans.

Table 3 below relates the preliminary results for a series of employment regressions with functional form of model (3). For these models, an approximation for hourly wage was created by multiplying the total number of hours worked per week by the total number of weeks worked per year and dividing yearly earnings by that number. This number is literally interpreted as earnings/hour or wage, but can be thought of as a good proxy for productivity in general. The “male” dummy below takes value one if the individual is male. The results are presented below, with all coefficients representing the percentage change in Rupiahs/hour. Finally, the initial analysis focuses on capturing the differences in wages for orphans in different age buckets without controlling for single or double orphanhood and without measuring gender differences, both of which will be addressed in later models.

Table 3: Effect of Orphanhood Status on ln(wage)

	(1) All Ages	(2) Age<14	(3) Age ≥14
Orphan	-0.247***	-0.042	-0.263***
	[2.45]	[0.24]	[2.30]
Male	0.465***	0.489***	0.400***
	[5.22]	[3.56]	[2.94]
Years of education	0.007	0.006	0.004
	[0.89]	[0.44]	[0.33]
Baseline age = 7	0.359	0.756	
	[1.50]	[3.10]	
Baseline age = 8	0.052	0.401	
	[0.21]	[1.58]	
Baseline age = 9	0.462	0.653	
	[2.19]	[3.04]	
Baseline age = 10	0.384	0.621	
	[3.02]	[5.41]	
Baseline age = 11	0.402	0.629	
	[2.82]	[8.42]	
Baseline age = 12	0.190	0.401	
	[1.08]	[3.82]	
Baseline age = 13	0.281	0.545	
	[1.87]	[4.53]	
Baseline age = 14	0.121		-0.131
	[0.89]		[0.99]
Baseline age = 15	0.356		0.131
	[1.81]		[0.76]
Baseline age = 16	0.315		0.071
	[2.64]		[0.66]
Baseline age = 17	0.579		0.303
	[3.76]		[2.45]
Baseline age = 18	0.237		(omitted)
	[2.03]		-
Pre-tsunami HH size	0.008	0.020	0.028
	[0.39]	[0.61]	[0.90]
Mother's education	0.038***	0.034*	0.050***
	[3.20]	[1.88]	[3.24]
Constant	7.800	7.485	7.919
	[41.40]	[21.26]	[29.96]
Observations	634	275	359
Number of kecamaten	58	53	56
R-squared	0.09	0.10	0.08

Robust t statistics in brackets
* p < .1, ** p < .05, *** p < .01

We find encouraging results that seem to agree with our education outcomes. Model one is a standard fixed effects model controlling for both gender and age. The coefficient for "Orphan" (which takes the value 1 for both single and double orphans, 0 for non-orphans) is negative and statistically significant. However, when measuring only those respondents younger than 14, the orphan parameter becomes insignificant. The covariate retains significance when controlling for age 14 and older. For all ages, orphans earn 25% less than non-orphans. Upon controlling for age, adolescent orphans show earnings 26% less than non-orphans. This wage gap is large enough to alone suggest that there is a meaningful difference in quality of life between orphans and non-orphans. It should be noted that the large and significant "Male" covariate here and in the following tables is a result of the gender gap existing in Indonesian culture; a proper analysis and discussion of this gap is outside the scope of this paper, but should be kept in mind when we move on to studying gender differences in wages among orphans.

Finally, we find that years of completed education have a positive but small and insignificant effect on wages. For all models, an increase in one year of completed education increases wages by less than 1%. These results are inconsistent with existing economic theory, as we should expect education to have a tangible effect on wages. It is likely the case that the effect of education is being swept up by other parts of our model. First, note that age indicators are often positive and significant. While these variables will largely capture unobservable characteristics such as workforce experience, they also will retain some of the effects of education, as older children are more likely to have completed more years of education. Finally, note that our model is conditional on being employed. It is very likely that completed years of education primarily increase an individual's probability of being employed at all, and thus our models do not capture the significance of education directly.

These results suggest that orphans are less productive and earn less than their non-orphaned cohorts. The insignificant results for younger orphans may be explained in two ways: younger children may have not yet entered the workforce and would not have any wages to report (and are thus dropped from our regression, which is conditional on employment), or as we saw with our education results, younger orphans did not experience any significant changes with their education, which would be a significant predictor of wages later in life. However, older orphans are significantly adversely affected, again earning about 26% less than their non-orphan cohorts. Note that these models control for employment status so that any unemployed respondents were dropped. Thus, this large difference in wages cannot be explained by a larger portion of orphan respondents being unemployed. As such, Table 3 suggests that there exists a long-term affect of orphanhood that persists after schooling is completed and once the individual enters the workforce. The lower wages observed here are likely intimately linked to the lower levels of education among older orphans observed above. Further analysis is necessary before drawing any more conclusions.

Focusing on the older age group, a second set of regressions was run to attempt to measure the differences between male and female orphans. Table 4 contains these results.

Table 4: Effect of Orphanhood Status, Age ≥ 14, on ln(wage) with Gender Differences

	(1) Males	(2) Females
Orphan	-0.159	-0.565
	[1.18]	[1.58]
Years of education	-0.001	0.011
	[0.06]	[0.67]
Baseline age = 14	0.028	-0.161
	[0.16]	[0.57]
Baseline age = 15	0.066	0.322
	[0.32]	[0.86]
Baseline age = 16	0.123	0.157
	[0.88]	[0.61]
Baseline age = 17	0.297	0.436
	[1.65]	[1.59]
Baseline age = 18	(omitted)	(omitted)
	-	-
Pre-tsunami HH size	0.038	0.033
	[0.95]	[0.47]
Mother's education	0.019	0.100***
	[0.90]	[2.29]
Constant	8.533	7.327
	[17.56]	[13.61]
Observations	251	108
Number of kecamaten	55	43
R-squared	0.02	0.24

Robust t statistics in brackets
* p < .1, ** p < .05, *** p < .01

Model (1) measures the difference in log wages between aged 14 or older orphaned males and their non-orphaned cohorts, controlling for community fixed effects. Model (2) does the same for females. Note that statistical significance is lost when we attempt to measure a gender difference, which is more than likely a result of small sample size. Note that model (2) has a sample size of 108, which is significantly smaller than sample sizes above. Upon controlling for age, employment, and gender, we lose too many degrees of freedom to appropriately capture orphanhood effects. Future analyses in this vein would need to find a data set with a larger sample size in order to correct these issues. Nevertheless, note that the orphanhood effect for both genders is negative, which is consistent with our results from table 3. While the orphan effect on wages for females seems to be much larger than that of males, which runs contradictory to what we have seen thus far, it is encouraging to see that the difference between these coefficients is roughly .41, which is consistent with the "Male" coefficient from table 3. Thus, the wage-gender gap is clearly being swept into the male-female difference of table 4.

Finally, similar regressions were run to try to isolate the difference between single and double orphans. The results are presented in Table 5 with separate models estimated for maternal orphans, paternal orphans, and double orphans. As before, we control for community level fixed effects and estimate robust standard errors.

Table 5:	**Orphanhood Status on ln(wage) – Single vs Double Orphan**		
	(1) Maternal Orphan	(2) Paternal Orphan	(3) Double
Orphan	-0.170	-0.142	-0.372**
	[1.07]	[1.12]	[2.57]
Male	0.427***	0.437***	0.460***
	[4.47]	[4.62]	[4.97]
Years of education	0.006	0.010	0.007
	[0.75]	[1.25]	[0.92]
Baseline age = 7	0.334	0.370	0.326
	[1.35]	[1.63]	[1.35]
Baseline age = 8	0.037	0.051	0.053
	[0.13]	[0.22]	[0.21]
Baseline age = 9	0.516	0.504	0.463
	[2.41]	[2.61]	[2.24]
Baseline age = 10	0.425	0.407	0.365
	[3.05]	[3.58]	[2.77]
Baseline age = 11	0.447	0.429	0.388
	[3.03]	[3.63]	[2.71]
Baseline age = 12	0.203	0.159	0.177
	[1.15]	[1.05]	[1.03]
Baseline age = 13	0.302	0.297	0.264
	[1.96]	[2.07]*	[1.74]
Baseline age = 14	0.124	0.112	0.092
	[0.90]	[0.92]	[0.70]
Baseline age = 15	0.366	0.338	0.343
	[1.75]	[1.73]	[1.73]
Baseline age = 16	0.321	0.293	0.302
	[2.38]	[2.64]	[2.62]
Baseline age = 17	0.661	0.603	0.558
	[3.86]	[4.06]	[3.52]
Baseline age = 18	0.278	0.269	0.226
	[2.32]	[2.70]	[1.91]
Pre-tsunami HH size	0.008	0.012	0.006
	[0.35]	[0.53]	[0.27]
Mother's education	0.035***	0.036***	0.036***
	[2.72]	[2.70]	[2.96]
Constant	7.822	7.751	7.830
	[37.57]	[38.54]	[39.63]
Observations	580	581	634
Number of kecamaten	58	58	58
R-squared	0.08	0.08	0.09

Robust t statistics in brackets

* p < .1, ** p < .05, *** p < .01

We once again see negative results across the board. However, upon controlling for specific orphanhood status, note that the only statistically significant results are for double orphans, who are predicted to earn as much as 38% less than non-orphan cohorts. Again, these results seem to be corroborated by the narrative driven by our education analyses in which double orphans were adversely affected. These results suggest that the effects of orphanhood are long-term and undoubtedly bleed over into labor outcomes.

The results here, combined with those reported above, come together to form the beginning of a cohesive narrative on the effects of unexpected parental death. Orphanhood seems to disproportionately affect older male children, who likely would have to replace their deceased parent(s) as the new head of household and primary breadwinner. As a result, these young males would have to drop out of school early and take on relatively low-paying jobs. We have shown in our models that completed years of education is not a driving factor in the wage gap. Additionally, we should not expect that workforce experience can explain this difference either, as orphans would likely become employed earlier than non-orphans and thus have more experience, which would work in the opposite direction. Thus, it is likely the case that these orphans take on low-paying jobs out of necessity. They do not have time to shop around for the best job option, as they need to immediately supplement household income after losing a household head. Together, these results suggest that older, orphaned males experience what could be a called a forced arrival into adulthood. Interestingly, much of the prior literature focusing on AIDS related death states that it is young girls who may experience some sort of forced arrival into adulthood. Chae et al. (2013) and Bozzoli et al. (2016) both found that female orphans are more likely to have an early sexual debut and are more likely to have a teen pregnancy than non-orphan cohorts. Thus, we can continue to infer that there is a fundamental difference in how unexpected parental death, like that from a natural disaster, affects orphaned children.

VI. Conclusion and Future Work

Our work aims to discuss how orphanhood can have a long-term impact on various measures of socio-economic status. We have shown that older male orphans finish fewer years of education and earn less in the workforce, and that these effects are more pronounced for double orphans. There is no evidence that orphans recover within the 10-year period that STAR data covers. However, it is important to note that these results do not cover the entire scope of socio-economic well-being. As has been discussed, many pieces of prior literature will focus on fertility and marriage outcomes, particularly for orphaned girls. A brief analysis using STAR data produced the following results in table 6, with a logit model used to predict the probability that an individual is currently married. The coefficients here represent the change in probability that a respondent is currently married. Thus, a coefficient of -10 for the orphan variable would mean an orphan is 10% less likely to be currently married. All models are estimated with community level fixed effects and allow for robust standard errors. The models control for age in the same way as earlier models, but these indicators were dropped from the table for brevity.

Table 6:	Change in Percent Chance that Respondent is Married		
Variable	**Model 1 (Non-FE)**	**Model 2 (FE)**	**Model 3 (FE, Age >=14)**
Orphan	3.61	4.61	~
	[1.24]	[1.42]	~
Male orphan	~	~	14.33**
	~	~	[2.07]
Female orphan	~	~	-7.59
	~	~	[-1.03]
(1) if male	-19.93***	-20.05***	-44.20***
	[-10.87]	[-9.08]	[-5.94]
Age	~	~	~
	~	~	~

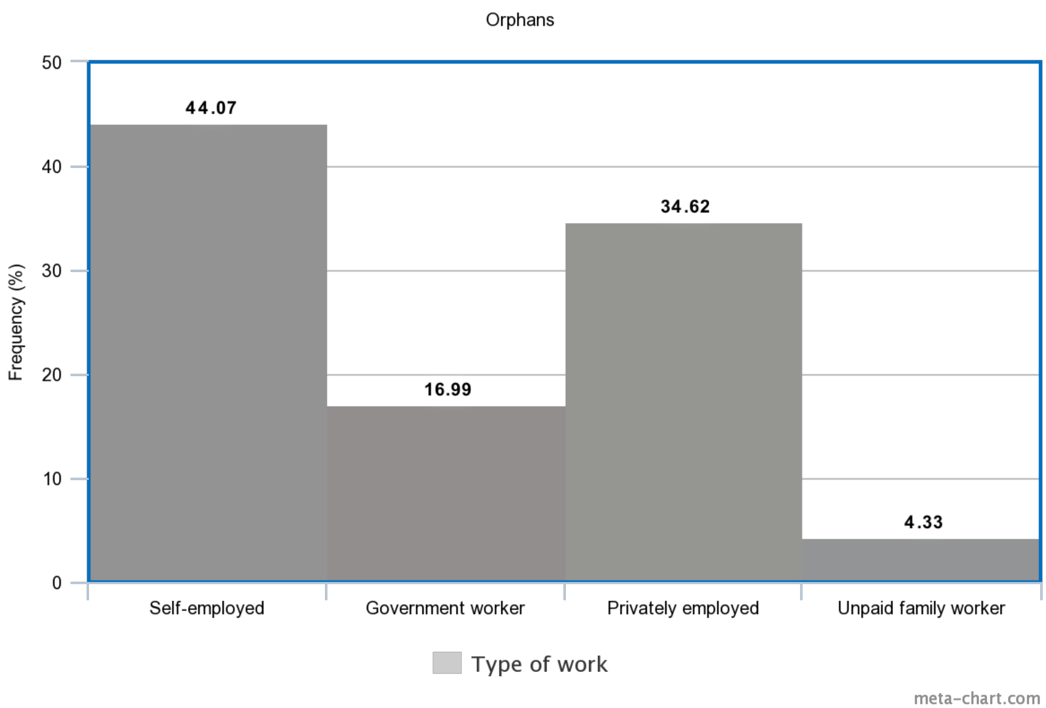
t statistics in brackets

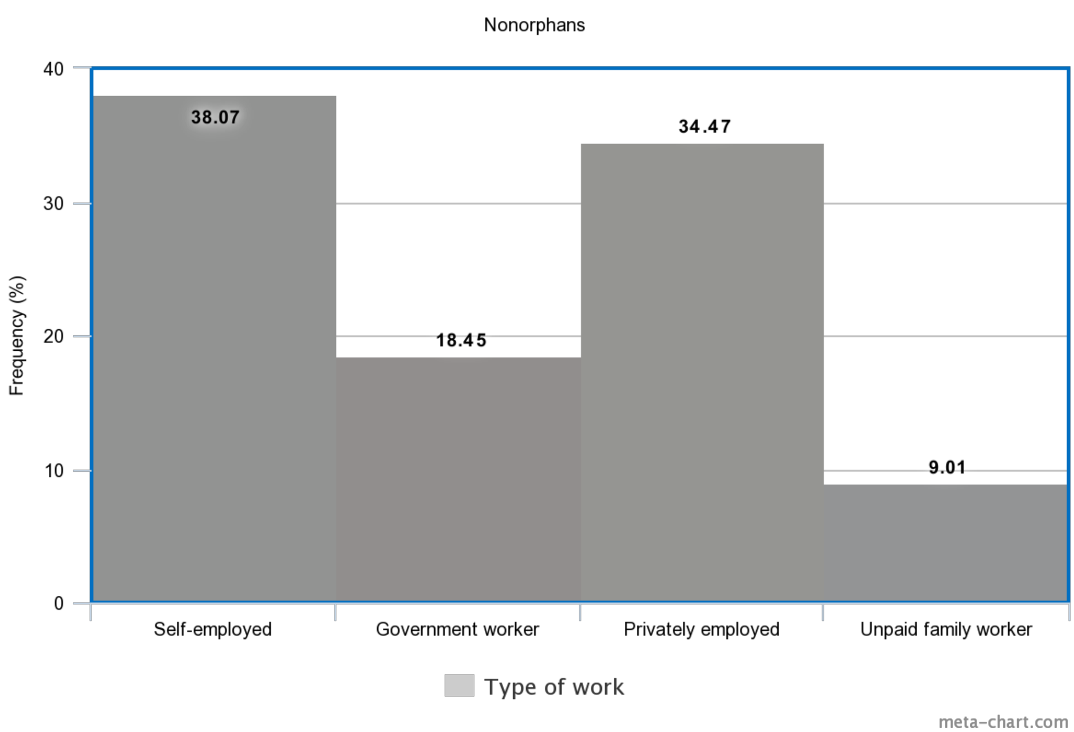
*** p<.01, **p<.05, *p<.1

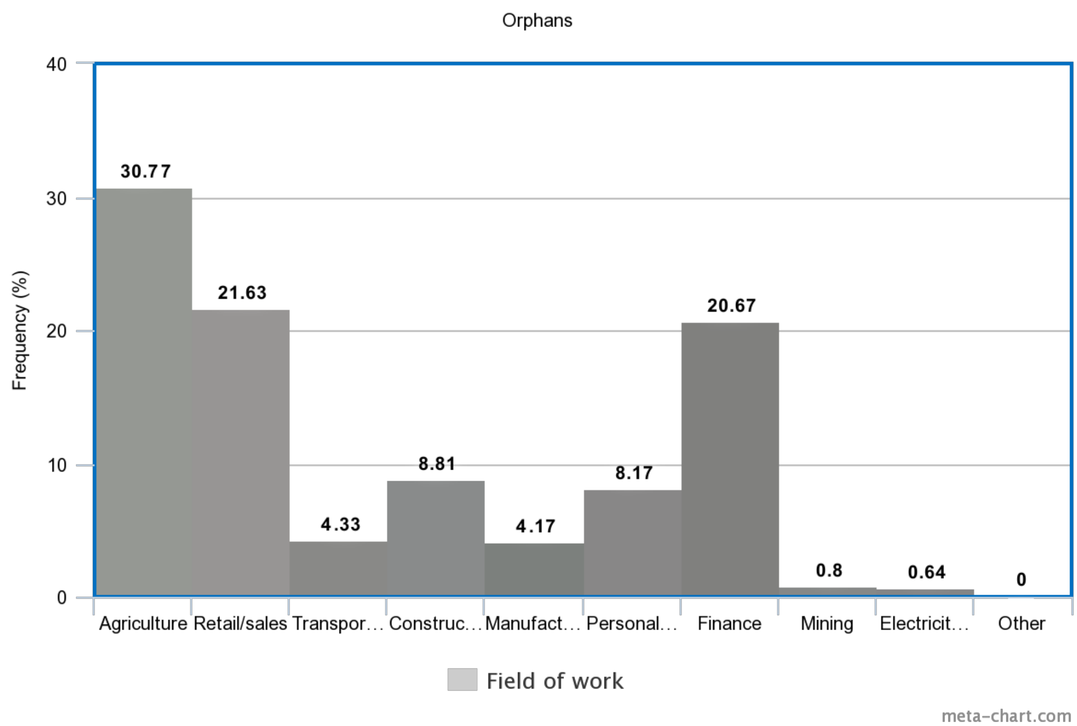
In efforts to avoid repetition, I will focus exclusively on the results stated under "Model 3". Here, a fixed-effect model is used to compare male and female orphans, who were aged 14 or older at the time of the tsunami, with their non-orphan cohorts. The only statistically significant result is that of older male orphans, who are 14% more likely to be married than their cohorts. Note that, for all three models, the "gender" dummy variable takes on large, negative, and statistically significant value. This is likely a result of the construction of the sample, which was given at the household level; relatively few households are solely headed by females in Indonesia, and thus we expect to find unmarried females to be under-represented here. Additionally, note that the youngest of our orphan sample would still be 16 at the time of our latest survey, which may still be too young to conduct a proper analysis of marriage outcomes. These complications leave this discussion outside the scope of this paper. However, there is valid reason to believe that marriage and fertility outcomes would be different between orphans and non-orphans. Thus, future literature in this vein may benefit from an in-depth discussion of marriage in order to complete the orphanhood narrative.

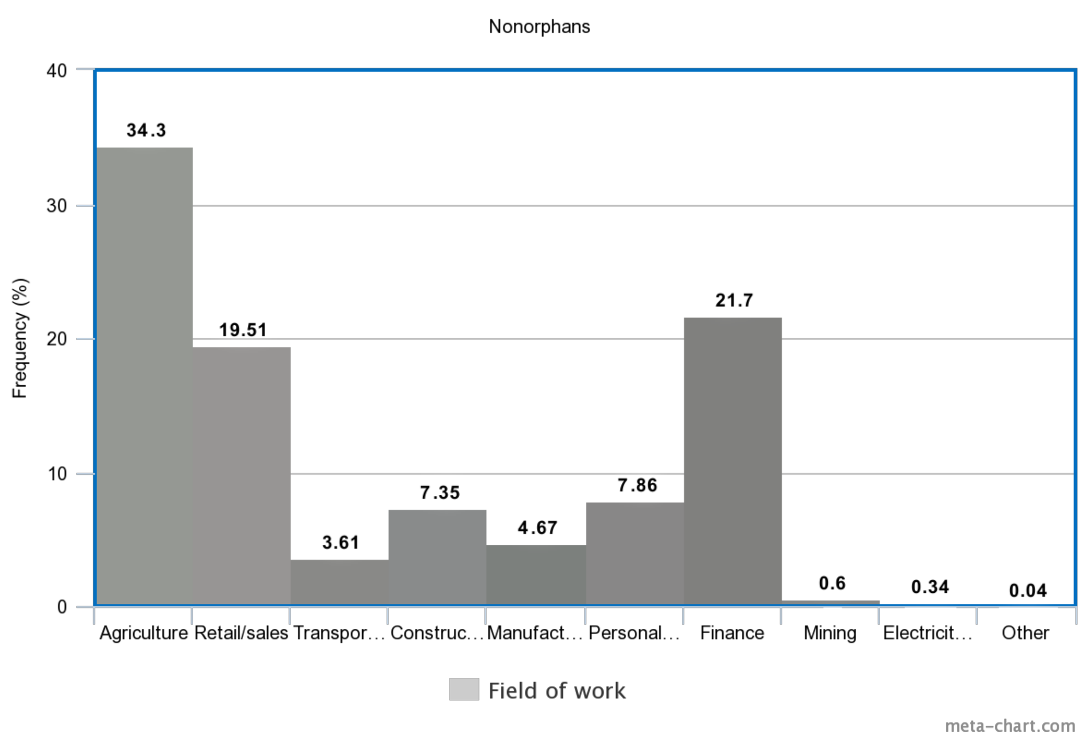
The orphan condition, as described by STAR data, is undeniably different from that described by literature focusing on AIDs related death in Sub-Saharan Africa. There are many factors at play that contribute to these differences; first, there exist myriad culture differences between Indonesia and Sub-Saharan Africa that could manifest themselves in our results. More importantly, however, is the unexpected nature of parental death within STAR data, which is unique to our analysis. There is no "period of erosion" to bias our results. Instead, we have a narrative in which young male orphans must step up as a stand-in head-of-household. They are forced to leave school early, and their labor market potential suffers as a result. Compare these effects to the narrative driven by Bozzoli (2016) in which younger, female orphans were pressured to supplement family income via transactional sex. How orphanhood affects individual children clearly varies between cultures as well as types of parental death. However, we expect that these adverse effects may continue throughout an individual's life. Do orphans keep they initial, low-paying jobs that they took on to help their households, or do they eventually mature into better paying work that more appropriately reflects their education background? As such, future literature should work to extend this analysis as far as possible.

Appendix A


Orphans
Frequency (%)
50
40
30
20
10
0
44.07
16.99
34.62
4.33
Self-employed
Government worker
Privately employed
Unpaid family worker
Type of work
meta-chart.com


Nonorphans
Frequency (%)
40
30
20
10
0
38.07
18.45
34.47
9.01
Self-employed
Government worker
Privately employed
Unpaid family worker
Type of work
meta-chart.com


Orphans
40
30
20
10
0
Frequency (%)
30.77
21.63
4.33
8.81
4.17
8.17
20.67
0.8
0.64
0
Agriculture
Retail/sales
Transpor…
Construc…
Manufact…
Personal…
Finance
Mining
Electricit…
Other
Field of work
meta-chart.com


Nonorphans
40
30
20
10
0
Frequency (%)
34.3
19.51
3.61
7.35
4.67
7.86
21.7
0.6
0.34
0.04
Agriculture
Retail/sales
Transpor…
Construc…
Manufact…
Personal…
Finance
Mining
Electricit…
Other
Field of work
meta-chart.com

References

Beegle, K., De Weerdt, J., Dercon, S. Orphanhood and the Long Run Impact on Children. *American Journal of Agricultural Economics*, 88(5), 1266-1272.

Berg, L., Rostila, M., Saarela, J., Hjern, A. Parental Death During Childhood and Subsequent School Performance. *Pediatrics*, 133(4).

Bicego, G., Rutstein, S., Johnson K. Dimensions of the emerging orphan crisis in sub-Saharan Africa. *Social Science and Medicine*, 56(6).

Bozzoli, Carlos. Orphanhood and fertility in young adults: Evidence from South Africa. *Economics and Human Biology*, 22, 190-200.

Cas, A.G., Frankenberg, E., Suriastini, W. et al. The Impact of Parental Death on Child Well-being: Evidence From the Indian Ocean Tsunami. *Demography,* 51(2).

Case, A.G., Paxson, C., Ableidinger, J. Orphans in Africa: Parental Death, Poverty, and School Enrollment. *Demography*, 43(3).

Chae, Sophia. Timing of Orphanhood, Early Sexual Debut, and Early Marriage in Four Sub-Saharan African Countries. *Studies in Family Planning,* (44(2), 123-146.

Evans, D. & Miguel, E. Orphans and Schooling in Africa: A Longitudinal Analysis. *Demography*, Volume 44(1), 35–57

Frankenberg, E., Gillespie, T., Preston, S. et al. Mortality, The Family and The Indian Ocean Tsunami. *The Economic Journal*, vol 121, issue 554, 2011.