Title

Essays on Women's Wellbeing in Developing Countries

Author

Lu, Frances

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays on Women's Wellbeing in Developing Countries

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Frances Rose Lu

Committee in charge:

Professor Prashant Bharadwaj, Co-Chair
Professor Tom Vogl, Co-Chair
Professor Titan Alon
Professor Gaurav Khanna
Professor Paul Niehaus

2024

The Dissertation of Frances Rose Lu is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

DEDICATION

To my parents, Xin Lu and Wei Feng, who have sacrificed so much to generate opportunities for me. I hope to pay it forward with my future career and endeavors.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This dissertation is the culmination of not only my own hard work but also my rewarding collaboration with mentors and peers.

First, I would like to thank my dissertation committee members. My committee co-chairs Prashant Bharadwaj and Tom Vogl greatly inspired both the topics I work on and the methods I work with. Paul Niehaus' mentorship both before and during my PhD helped me develop research and organizational skills that set me up for success. Titan Alon and Gaurav provided me with enthusiastic support and guidance throughout the research process.

Second, I would like to thank my wonderful friends and community in the department. In particular, I would like to thank the brilliant, strong women in my cohort whose support and encouragement allowed me to keep going when I could not find the strength within myself.

Finally, I would like to thank my wonderful family. Thank you to my parents, Xin Lu and Wei Feng, (and their parents) who sacrificed so much to provide opportunities for and improve the life chances of me and my siblings. This dissertation is dedicated to you. Thank you to my siblings, Amanda and Ben, who have always cheered me on.

Chapter 1, in full, is a reprint of the material as it appears in American Economic Review: Insights. Lu, F., & Vogl, T. (2023). Intergenerational persistence in child mortality. American Economic Review: Insights, 5(1), 93-109. The dissertation author was a primary investigator and co-author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in The Journal of Development Economics. Lu, F., Siddiqui, S., & Bharadwaj, P. (2021). Marriage outcomes of displaced women. Journal of Development Economics, 152, 102684. The dissertation author was a primary investigator and co-author of this paper.

VITA

EDUCATION

2016	Bachelor of Science in Engineering in Operations Research and Financial Engineering, Princeton University
2024	Doctor of Philosophy in Economics, University of California San Diego

PUBLICATIONS

Lu, F., Siddiqui, S., & Bharadwaj, P. (2021). Marriage outcomes of displaced women. Journal of Development Economics, 152, 102684.

Lu, F., & Vogl, T. (2023). Intergenerational persistence in child mortality. American Economic Review: Insights, 5(1), 93-109.

FIELDS OF STUDY

Development and Labor Economics

ABSTRACT OF THE DISSERTATION

Essays on Women's Wellbeing in Developing Countries

by

Frances Rose Lu

Doctor of Philosophy in Economics

University of California San Diego, 2024

Professor Prashant Bharadwaj, Co-Chair
Professor Tom Vogl, Co-Chair

In Chapter 1, we study the intergenerational persistence of inequality by estimating grandmother-mother associations in the loss of a child, using pooled data from 119 Demographic and Health Surveys in 44 developing countries. Compared with compatriots of the same age, women with at least one sibling who died in childhood face 39% higher odds of having experienced at least one own-child death, or 7 percentage points at age 49. Place fixed effects reduce estimated mortality persistence by 47%; socioeconomic covariates explain far less. Within countries over time, persistence falls with aggregate child mortality, so that mortality decline disproportionately benefits high-mortality lineages.

In Chapter 2, we examine the marriage market outcomes of forcibly displaced women. Using data from 12 representative surveys in 7 countries, we document that women who are adolescents at the time of displacement are more likely to be married. This pattern is robust to the choice of control group and across countries. We do not find this pattern for displaced adolescent men. We provide additional evidence of this relationship by using unique features of the partition of India in 1947, an event that resulted in large-scale bilateral displacement between India and the newly formed Pakistan. Using a representative household survey collected in 1973, we find that women who were adolescents when they were displaced by partition were significantly more likely to marry earlier, in line with the descriptive cross-country evidence.

In Chapter 3, we estimate mother-daughter associations in IPV victimization using representative survey data from 16 countries in Sub-Saharan Africa, a region of the world with particularly high IPV rates. We find that persistence is large in both levels and relative magnitudes. Women who report that their mother was physically abused by their father are 26 percentage points (1.9x) more likely to report physical or sexual violence from their current spouse. We then explore the role of cultural transmission. We find robust evidence for parental socialization, where daughters' exposure to interparental violence increases their reported acceptability of IPV, and mixed evidence for the role of assortative matching in the marriage market on these attitudes.

Chapter 1

Intergenerational Persistence in Child Mortality

Classic theoretical models of economic development study whether inequality persists across generations and how it interacts with the process of aggregate growth [Banerjee and Newman, 1993, Galor and Zeira, 1993]. However, the scarcity of multigenerational data on income, consumption, and wealth in developing countries has hindered efforts to characterize intergenerational mobility.[1] Using multigenerational data from 44 developing countries, we estimate intergenerational persistence in a separate determinant of wellbeing—the death of a child—and study how it changes with aggregate progress against child mortality.

We rely on survey data to estimate associations between grandmothers and mothers in the loss of a child. Our analysis hinges on the coexistence, in a single survey, of a sibling history module—which asks a woman for survival information on all of her mother's children—and a birth history module—which asks for the same on her own children. In such a survey, one can estimate mortality persistence as the relationship between sibling and own-child mortality. To this end, we assemble the 119 Demographic and Health Surveys with both modules, providing data on 2.6 million births to 1.3 million women in contexts spanning varying levels of socioeconomic development and varying stages of the mortality transition.

[1] Recent work focuses on educational mobility, using survey reports of own and parental education or census data on coresident parents and children [Torche, 2014, Neidhöfer et al., 2018, Narayan et al., 2018, Asher et al., 2020, Alesina et al., 2021, Muñoz, 2021].

The data reveal significant intergenerational persistence in the loss of a child. Within a country at a given age, women with at least one sibling who died under 5 face 39% higher odds of losing at least one child under 5. In absolute terms, the risk gap accumulates to 7 percentage points for women in their late 40s. At the child level, we find that the odds of dying under 5 rise 9% with each additional maternal sibling under-5 death, or 1 percentage point of risk. We refer to percentage changes in odds as *proportional* persistence and percentage point changes in risk as *absolute* persistence.

The interpretation of these large magnitudes depends on one's exact interest in mortality persistence. If one cares about the intergenerational persistence of life chances, then persistent mortality *risk* is more relevant than persistent mortality *outcomes*. The same is true if one views child mortality as a proxy for other forms of family disadvantage. The observed death of a sibling is an inherently noisy proxy for mortality risk or the socioeconomic determinants of mortality. This noise implies that our estimated outcome association understates the underlying risk association.[2] However, if one cares about the joint experience of sibling and maternal bereavement, then the outcome association is directly of interest.

Child mortality may persist across generations because of genetic inheritance, cultural transmission, socioeconomic immobility, the long reach of maternal childhood health, or geographic health disparities. To shed light on some of these channels, we conduct a simple accounting exercise that can be valuable for future work that seeks to identify causal mechanisms, analogous to how development accounting informs research into the mechanisms of economic development [Caselli, 2005]. The data allow us to assess the last three channels by adding covariates and fixed effects to the mortality persistence regression. We find that place fixed effects account for 47% of mortality persistence, while covariates for maternal human capital, wealth, and maternal health account for far less. The fixed effects are based on place of residence, so in principle, they may reflect spatial sorting rather than place effects *per se*. However, we

[2] Vaupel [1988] makes the similar point that the intergenerational association of lifespan understates the intergenerational association of relative mortality hazards.

find that place fixed effects also halve mortality persistence in a subsample of women who have lived in the same place all their lives, casting doubt on the sorting explanation. More likely, the place fixed effects reflect the persistence of place—including disease ecology, public health infrastructure, and health care access—or perhaps unmeasured dimensions of human capital and wealth. Separate from these channels, we also find that neither mothers' nor grandmothers' fertility mediates persistence in cumulative child deaths experienced over the lifecycle.

Aggregate mortality decline may magnify or reduce mortality persistence, depending on its distribution across high- and low-mortality lineages. This link has close analogies in the Kuznets Curve [Kuznets, 1955], which describes how income inequality changes with income growth, and the Great Gatsby Curve [Krueger, 2012], which describes how intergenerational income mobility relates to income inequality. To assess it, we estimate mortality persistence separately by country and the child's five-year birth period. We then regress the estimates on UN under-5 mortality rates, with country and birth period fixed effects. We find that as aggregate mortality declines within a country over time, absolute mortality persistence falls, but proportional mortality persistence does not. Our results suggest that the decline of child mortality in the late 20^{th} and early 21^{st} centuries had greater absolute benefits for lineages with historically higher child mortality, but not enough to close their relative disadvantage.

Analysts of sibling history data have long worried that respondents omit deceased siblings [Helleringer et al., 2014]. Based on independent responses from coresident mothers and daughters, we estimate that respondents underreport sibling childhood deaths by 11%. However, we find in Monte Carlo simulations that even if sample-wide underreporting were twice as extensive, bias in our estimators of mortality persistence would be small.

Our research contributes to a multidisciplinary literature on intergenerational associations in mortality and health. Biodemographers and behavioral geneticists have long studied family-level variation in longevity [Cohen, 1964, Vaupel, 1988, Herskind et al., 1996, Iachine et al., 2006], finding moderate heritability. Economists have until recently focused on early-life health, finding associations between mothers' height and children's health in poor countries

[Venkataramani, 2011, Bhalotra and Rawlings, 2013], and between mothers' and children's birth weights' in rich countries [Black et al., 2007, Currie and Moretti, 2007, Royer, 2009].[3] More recent work in economics includes adult morbidity [Halliday et al., 2021] and longevity [Black et al., 2022] in rich countries. Scaled appropriately, our estimates of child mortality persistence are comparable to US intergenerational associations in birth weight and lifespan but smaller than that in adult morbidity.

The intergenerational persistence of child mortality also relates to two literatures on the welfare consequences and policy implications of multidimensional inequality. Recent work builds quantitative tools to capture the contribution of mortality to cross-country welfare inequality [Becker et al., 2005, Fleurbaey, 2009, Fleurbaey and Gaulier, 2009, Jones and Klenow, 2016] and the prevalence of deprivation [Baland et al., 2021]. Our results endorse extending these tools to study persistent within-country mortality inequality. An older literature documents that mortality can fall without economic growth, using China, Cuba, Kerala, and Sri Lanka as examples [Caldwell, 1986, Sen, 1999]. Their experiences suggest that the policy options for combating persistent inequality may go beyond conventional tools like progressive taxation and redistribution. Public health programs, for example, may reduce geographic mortality dispersion and thus reduce persistent inequality in life chances. Bhalotra and Rawlings [2013] find that maternal height becomes less related to infant mortality as immunization rates rise. Yet they, like us, find little role for long-run income growth.

1.1 Data

We draw on the Demographic and Health Surveys (DHS), which interview women of childbearing age (15-49) [ICF, 2004-2017]. The surveys include a birth history module, which logs the respondent's live births and their survival. A subset also includes a sibling history module, which logs the respondent's reports of her mother's live births and their survival.

[3]The birth weight studies find that early-life conditions rather than genetics account for much of the intergenerational correlation.

We use the 119 surveys with both modules, leading to a 44-country sample. The countries are concentrated in sub-Saharan Africa but are also scattered across other developing regions (Appendix Table A.1).

The surveys have data on 1.7 million women, but our analysis necessitates two major sample restrictions. First, we omit respondents without siblings because they are uninformative about mortality persistence. Second, we omit respondents under 20 years old because their sibships may be incomplete, and they are unlikely to have given birth at least 5 years before the survey. We apply two minor restrictions, omitting respondents missing key covariates and those older than the standard DHS maximum age of 49. These restrictions leave 1.3 million women aged 20-49. When we analyze birth-level data, we exclude births missing under-5 mortality status or occurring more than 20 years before the survey, leaving 2.6 million births.

Descriptive statistics on the 1.3 million women appear in Table 1.1. On average, respondents are 32 years old and have 4.9 years of education. Roughly two-thirds live in rural areas. Respondents have fewer children than their mothers: children ever born averages 3.4, while siblings ever born averages 5.7. This discrepancy reflects both fertility decline and incomplete childbearing for younger respondents.

The sibling histories detect substantial mortality, with 32% of respondents reporting at least one sibling under-5 death. Surprisingly, however, the birth and sibling histories show similar own-child mortality rates for respondents and their mothers. Respondents averaged 0.43 dead before age 5, their mothers 0.61. Dividing by children or siblings ever born, both imply under-5 mortality risk of 11-13% per birth. Given the decline of infant and child mortality since the mid-20th century, the similarity may suggest that respondents underreport deceased siblings. An alternative explanation is that our sample of adult respondents necessarily overweights families with more surviving children.[4] We discuss reporting errors and their consequences in Section 1.3.2.

[4]If child mortality risk were independent of sibship size, this downward bias would exactly counterbalance an upward bias from omitting the respondent, who survived childhood [Trussell and Rodriguez, 1990]. In our data, however, sibling under-5 mortality rises with sibship size (Appendix Figure A.1).

Figure 1.1 shows basic patterns over the lifecycle. Older age groups exhibit higher risk of any child death and a larger gap by sibling mortality status, reflecting the accumulation of exposure with age and higher child mortality in earlier cohorts. At 20-24, 13% of women with deceased siblings have experienced child loss, compared with 9% among women without. By 45-49, these shares grow to 52% and 40%, respectively; the absolute gap more than doubles. These gaps partly reflect cross-country and cross-cohort variation, as women with deceased siblings tend to be from high-mortality countries and cohorts. The shaded area isolates the weighted average of *within-survey* gaps, accounting for 65-70% of the overall gap. Comparing women of the same age in the same survey removes variation between countries and cohorts.

The widening absolute gap suggests a proportional model. Appendix Figure A.2 rescales the vertical axis to log odds, revealing constant proportional gaps across age groups, with 56-63% higher odds for women with at least one sibling death, or 36-41% within survey. These results suggest that a proportional model will best capture the accumulation of differential child mortality risk over the lifecycle.

1.2 Methods

We estimate proportional models to accommodate lifecycle variation. At the woman level, we specify a generalized linear model of the form:

$$\eta_{js} = \gamma_s + \sum_a \alpha_a age^a_{js} + \beta_1 dead_{js} + \beta_2 sibs_{js} \tag{1.1}$$

for woman j in survey s. The response η_{js} is a function of *child* deaths, while the covariate $dead_{ms}$ is a measure of *sibling* deaths. To shed light on the role of grandmothers' fertility, we report results with and without controlling for the number of siblings ever born, $sibs_{js}$. For an exclusive focus on within-survey inequality, we include a survey-specific intercept γ_s. Finally, because risk cumulates over the lifecycle, we control flexibly for age by including single-year age indicators. We restrict age effects to be the same for all surveys to reduce computational burden but show in the Appendix that our main results do not change when we allow them to vary by survey.[5]

For our headline result, we estimate a logit regression relating the occurrence of any child death to the occurrence of any sibling death at the woman level. In terms of equation (1.1), η_{js} is the log odds of the mother experiencing at least one child death, and $dead_{js}$ is an indicator for the grandmother experiencing at least one child death. Of the specifications we run, this one has the clearest interpretation. It also allows us to predict the sign of any bias from underreporting of sibling deaths, since non-differential underreporting will attenuate the coefficient on the binary version of $dead_{js}$.

To understand the variation driving our main results, we estimate several model variants. First, to flexibly accommodate the changing distribution of sibling deaths, we also report estimations in which $dead_{js}$ is the count of sibling under-5 deaths. Second, for insight into the role of the respondent's fertility, we compare a Poisson regression of the count of child deaths at the woman level with a logit regression of mortality at the birth level. If women from high-mortality lineages accumulate more own-child deaths because they have more children, then the estimated parameters from the woman-level Poisson model will be larger than those from the birth-level logit model. The Poisson model is another version of equation (1.1), in which η_{js} is the logarithm of the expected number of child deaths experienced by the woman. However, the birth-level logit model requires a new specification:

$$\eta_{ijs} = \gamma_s + \beta_1 dead_{js} + \beta_2 sibs_{js} \tag{1.2}$$

where η_{ijs} now refers to the log odds of under-5 death for birth i to mother j in survey s. We omit the mother's current age because it is no longer directly related to cumulation of risk.

These functional forms treat mortality persistence as proportional, such that the odds

[5] The flexible specification effectively controls for country-specific cohort variation, since age and cohort are collinear conditional on survey year.

or expected count of child deaths are proportional to the occurrence or count of sibling deaths. Our main estimands are odds ratios and incidence rate ratios. Because absolute persistence is of independent interest, we also compute average marginal effects. All analyses of individual-level data use sampling weights rescaled to reflect each survey's contribution to the sample and cluster standard errors at the survey cluster (village or city block) level.

1.3 Results

1.3.1 Pooled Estimates

Pooled estimates of child mortality persistence appear in Table 1.2. The top of each panel reports exponentiated coefficients: odds ratios for the logit models and incidence rate ratios for the Poisson models. Average marginal effects appear at the bottom.

Panel A uses an indicator for at least one sibling under 5 death as the main covariate. Column (1) indicates that having at least one deceased sibling is associated with a 39% increase in the odds of having at least one deceased child. To capture the cumulation of risk over the lifecycle, we compute the average marginal effect at the last age in the sample, 49. Women with deceased siblings are 7 percentage points more likely to have experienced a child death by the end of reproductive age. Our proportional model fits the cumulation of differential mortality risk over a woman's reproductive years well. Appendix Figure A.3 reports similar results in pooled and age-specific estimations, with odds ratios stable at roughly 1.4 over the lifecycle, and average marginal effects roughly doubling from the early-20s to the late-40s. Stable odds ratios imply expanding absolute gaps.

The remainder of Table 1.2 compares alternative models of child mortality persistence, with results suggesting that neither the woman's fertility nor her mother's explains child mortality persistence. All six models—the woman-level logit, the woman-level Poisson, and the birth-level logit, with the sibling death indicator (Panel A) or count (Panel B)—find significant persistence. Within each model, the estimate changes little when we include siblings ever born as a covariate, suggesting that grandmothers' fertility plays little role in the results. A comparison of the woman-level Poisson estimates (columns [3]-[4]) to the birth-level logit estimates (columns [5]-[6]) further suggests that mothers' fertility plays little role. The incidence rate ratios from the woman-level models are similar to the odds ratios from the birth-level models, implying that risk to a woman cumulates in proportion to her number of children. If mortality persistence partly operated through a correlation of family mortality risk with fertility, then the woman-level estimates would exceed the birth-level estimates.

Of these remaining estimates, that in Panel B, column (6) is most intuitive: how an additional sibling death relates to the odds of death for a given child, holding fixed the number of siblings. This specification is appealing because it holds family size fixed in both generations. It finds that each additional under-5 death of a mother's sibling is associated with a 9% increase in the odds of an own-child's death, or a 1 percentage point increase in the probability of death.

1.3.2 Interpretation and Robustness

Our results raise three concerns: about the magnitudes of the estimates, the conflation of intra- with inter-cohort variation, and bias from underreporting of sibling deaths.

Magnitudes

The results in Table 1.2 are straightforward to interpret by themselves. But how does sibling death compare to other determinants of child mortality, and how does child mortality persistence compare with other intergenerational associations in health?

In our data, socioeconomic differentials in any child death somewhat exceed cumulative mortality persistence. Appendix Figure A.4 reports that the odds ratio for rural residence is 2.05, and that for not finishing primary school is 2.65, compared to the mortality persistence odds ratio of 1.39. To find a mortality persistence odds ratio above 2, one has to look to women from extremely high mortality families. Relative to women with no deceased siblings, women with eight or more have 2.29 the odds of losing at least one child. However, such women comprise less than 0.2% of the sample.

When we compare child mortality persistence to other intergenerational health associations in the US, we find that it is similar to the birth weight and lifespan associations but smaller than the adult morbidity association. Currie and Moretti [2007] find that children with low birth weight mothers are 3.9 percentage points more likely be low birth weight themselves, implying an odds ratio of 1.9. We find a larger marginal effect but smaller odds ratio. To obtain quantities that we can compare to intergenerational associations in lifespan and adult morbidity, we compute partial correlations of sibling and own-child death, conditional on survey indicators (Appendix Table A.2).[6] The correlation in any death is 0.08, and the correlation in the count of deaths is 0.07. These magnitudes are similar to Black et al.'s (2022) estimate of 0.09 for the mother-daughter correlation in age at death and somewhat smaller than Halliday et al.'s (2021) estimates of intergenerational associations in adult morbidity (proxied by self-reports of health status and health conditions), which imply mother-daughter correlations of 0.11-0.17.

Cohort Effects

Our estimations include survey indicators but not cohort indicators, so they mix within-cohort comparisons of women from low- and high-mortality families with between-cohort comparisons of women from low- and high-mortality cohorts. Our woman-level models include age indicators, but because the data were collected in a variety of times and places, the age indicators do not fully absorb sample-wide cohort variation, let alone country-specific cohort variation. Appendix Figure A.5 adds interactions of survey indicators and age group indicators to all 12 models in Table 1.2, finding no change in the estimates.[7] The survey-by-age group effects effectively control for country-specific cohort variation, so our results do not reflect comparisons of cohorts with high and low sibling mortality. Instead, they are primarily driven by within-cohort comparisons.

[6]To avoid issues of proportionality over the lifecycle, we limit this analysis to women aged 45-49.

[7]We use five-year rather than one-year age groups in these interactions because the latter proved computationally burdensome.

Reporting Errors

Respondents may underreport the childhood deaths of their siblings. To assess the extent of underreporting, we follow Masquelier and Dutreuilh [2014] in studying coresident mothers and daughters. The DHS instructs enumerators to interview each 15-49 year old female household member alone, so a mother's birth history provides an independent check on her daughter's sibling history. Linkage of mother and daughter respondents is possible in 85 of the 119 surveys; in these surveys, 38% of 15-19 year olds coreside with their 30-49 year old mothers. In these mother-daughter pairs, daughters' reports of sibling under-5 deaths are highly correlated with mothers' reports of child under-5 deaths ($\rho = 0.90$), with moderate underreporting and minimal overreporting (Appendix Tables A.3-A.4). When mothers report no child deaths, 98% of daughters report no sibling deaths. When mothers report at least one child death, 86% of daughters report at least one sibling death, with the exact counts matching in 77-80% of cases. Overall, daughters report 11% fewer deaths than mothers.

Bias from underreporting depends on whether we use the count of sibling deaths or an indicator for any sibling death. For the indicator, if underreporting is nondifferential, then it biases us toward finding no persistence [Davidov et al., 2003, Mahajan, 2006].[8] For the count, the bias is more difficult to characterize analytically [Bound et al., 2001]. Due to these ambiguities, we perform Monte Carlo simulations of a plausible form of measurement error. We simulate the omission of deceased siblings assuming a fixed probability of omission per deceased sibling. After each draw, we drop observations with 0 reported siblings and estimate the 12 regressions in Table 1.2. We assess how the estimates change as we increase the probability of omission from 0 to 25% (over twice the rate in the mother-daughter pairs).

Even at large probabilities of omission, the simulated biases in the mortality persistence estimates are small (Appendix Figure A.6). For the mother-level regression of any child death on any sibling death, a 25% probability of omission leads to a mean odds ratio of 1.378, compared

[8]However, women who erroneously report having no siblings drop out of the sample, which may have different consequences from misclassification alone.

to our original result of 1.386. For the birth-level regression of child death on the number of sibling deaths and siblings ever born, a 25% probability of omission leads to a mean odds ratio of 1.105, compared to our original result of 1.092. The simulations suggest that our estimations are robust to underreporting.

1.3.3 Accounting for Mortality Persistence

What channels account for intergenerational persistence in child mortality, and how does it vary within and between populations? We assess channels by adding covariates, and we explore heterogeneity by estimating persistence separately by country and gender. The analysis of gender heterogeneity requires birth-level data, so we use the birth-level logit specification from Table 1.2, Panel B, column (6) throughout.

Adding covariates

Potential explanations for intergenerational persistence in child mortality include socioeconomic immobility, geographic health inequality, cultural transmission, health transmission, and genetic inheritance. To shed light on some of these explanations, Table 1.3 adds covariates and fixed effects to the regression. Our analysis is limited by the scope of the DHS questionnaire, which collects no genetic information and asks only a handful of questions on socioeconomic, geographic, and adult health outcomes. We account for place by including survey cluster (village or city block) fixed effects in a conditional logit model. Because the conditional logit model requires outcome variation within all units, we omit clusters lacking variation in child mortality from the analysis. For socioeconomic status, we use the mother's years of education and the DHS household wealth index, formed by taking the first principal component of a vector of indicators for durable goods ownership and improved housing conditions at the time of the interview. For maternal health, we use the mother's height. This accounting exercise is useful for pinpointing margins of interest for future research but not for directly identifying causal mechanisms.

Panel A finds that geography accounts for much of the intergenerational relationship, while measured socioeconomic status accounts for fall less. Column (1), which adds no further covariates or fixed effects, finds an odds ratio of 1.085, similar to Table 1.2, Panel B, column (6). Column (2) adds survey cluster fixed effects, shrinking the odds ratio to 1.043, a 49% reduction toward unity. Further adding the mother's education and the household's wealth index (column [3]) results in an odds ratio of 1.039, only 14% smaller. Thus, women with siblings who died in childhood live in places with higher under-5 mortality rates, which explains nearly half of the increased mortality odds their own children face. Place may capture disease ecology, public health infrastructure, and health care access, but also omitted socioeconomic variables. However, measured socioeconomic variables account for a limited share of the remaining persistence of child mortality.

Place contributes to health inequality [Burstein et al., 2019] and socioeconomic immobility [Asher et al., 2020, Alesina et al., 2021, Muñoz, 2021], so its primacy in the intergenerational persistence of child mortality is not surprising. Nevertheless, because the fixed-effect models condition on *current* place, whether unhealthy places are inherited in childhood or attained in adulthood through residential sorting is unclear. The DHS has limited information on childhood place of residence, but 77 of the 119 surveys in the sample ask whether the respondent has lived in her current place of residence all her life. To partially assess the role of residential sorting, we compare estimates with and without cluster fixed effects using data on only non-migrants, who constitute 44% of the surveys with information on migrant status. The sample restriction does not entirely solve the interpretation issue, since the non-migrant subsample is self-selected, but similar results in the non-migrant subsample and the full sample would suggest a role for the inheritance of place in childhood.

Panel B analyzes non-migrants from the 77 surveys with information on migrant status, finding little change from the full sample results. Among non-migrants living in survey clusters with variation in child death, the addition of cluster fixed effects shrinks the odds ratio from 1.073 to 1.041, so that place accounts for 44% of mortality persistence. Given the resemblance to the full-sample result in Panel A, these findings suggest that the inheritance of unhealthy places accounts for half of the intergenerational persistence of child mortality.

To consider the role of maternal health, Panel C controls for the respondent's height in the 103 surveys with relevant data, finding that it explains little of the mortality association. Adult height is a proxy for early-life health health [Currie and Vogl, 2013] and the basis for Bhalotra and Rawlings' (2013) research on intergenerational health transmission. Although height negatively predicts child survival, it does not change the odds ratio on deceased siblings.[9]

Heterogeneity

Appendix Figure A.7 reports birth-level odds ratios by country. Estimates are uniformly greater than 1, but with considerable heterogeneity, ranging from 1.02 to 1.24. Appendix Table A.5 investigates gender heterogeneity in the pooled birth-level model, estimating separate regressions for boys and girls, as well as splitting the count of sibling deaths into brother and sister deaths. We find little heterogeneity, with odds ratios of 1.08-1.11 for brothers, sisters, boys, and girls.

1.3.4 Mortality Persistence over the Mortality Transition

The country-level heterogeneity in Appendix Figure A.7 raises the question of how the intergenerational persistence of child mortality varies with aggregate health conditions. The answer sheds light on the progressivity of mortality decline: whether previously high- or low-mortality lineages benefit more from aggregate health improvement. To investigate, we relate mortality persistence to the aggregate under-5 mortality rate. We consider mortality conditions at the time of each *child's* birth, as data on conditions at the time of the *mother's* birth are not uniformly available and are more subject to recall error.

The results of this exercise may depend on whether we measure mortality persistence in proportional or absolute terms. The relationship between between the odds ratio and the marginal effect changes as mortality falls in aggregate. Appendix Figure A.8 demonstrates this point by

[9] Consistent with Bhalotra and Rawlings [2013], the conditional logit finds that a 10 centimeter increase in height predicts 10% lower odds, or a 1.4 percentage point lower probability, of under-5 death.

drawing the relationship between the marginal effect and the odds ratio for a binary covariate.[10] Holding baseline mortality fixed, the marginal effect increases with the odds ratio; holding the odds ratio fixed, the marginal effect increases with baseline mortality. At a higher mortality rate, a given proportional association implies a larger absolute association.

This three-way linkage of the aggregate mortality rate, the odds ratio, and the marginal effect implies that proportional mortality persistence and absolute mortality persistence need not move in the same direction during aggregate mortality decline. Both could decrease, both could increase, or the marginal effect could decrease while the odds ratio increases. The only impossibility is for mortality decline to be accompanied by a (weakly) rising marginal effect and a (weakly) falling odds ratio.

We assess which of these scenarios best characterizes child mortality decline in the developing world in the late 20^{th} and early 21^{st} centuries. Because we are interested in mortality decline, a within-country phenomenon, we construct a country-period panel of persistence estimates and merge it with UN country-period estimates of child mortality rates [United Nations and Social Affairs, 2019]. To form this panel, we run the birth-level logit regression from Table 1.2, Panel B, column (6)—of child death on the number of deceased and ever-born siblings, conditional on survey indicators—for each country by five-year birth period cell.[11] The five-year periods correspond to when the respondent gave birth, not when she was born.[12] Just as in the earlier birth-level analyses, each respondent may enter the sample multiple times. The analysis seeks to describe how the gains from mortality decline are distributed across women giving birth in a particular period.

We then estimate linear regressions of these estimated parameters on the contemporaneous under-5 mortality rate (from the UN), country fixed effects, and period fixed effects. In some specifications, we additionally control for the log of real annualized GDP per capita from the

[10]If the covariate is multivalued, the relationship depends on its distribution.

[11]The birth-level data allow us to ask whether persistence parameters vary by year, and the count of sibling deaths captures the changing distribution of sibling mortality as mortality falls.

[12]UN mortality data are available for all periods in which respondents gave birth but not for all periods in which respondents were born.

Penn World Table [Feenstra et al., 2015] and an indicator for armed conflict from UCDP/PRIO [Gleditsch et al., 2002, Pettersson et al., 2021]. The coefficient on the under-5 mortality rate represents how mortality persistence changes with overall mortality within a country over time. We cluster standard errors by country to account for within-country serial correlation.

Country-level time series of the under-5 mortality rate highlight two points (Appendix Figure A.9). First, the Rwandan genocide was a singular mortality event, with an under-5 mortality rate over 40% higher than the next highest. Because we are interested in secular decline rather than shocks, and because the extent of mortality during the genocide is extremely uncertain, we omit this episode from our sample.[13] Second, under-5 mortality predominantly trended downward in sample countries over the sample period. This result motivates our interest in the distribution of mortality decline.

Table 1.4 reports the two-way fixed effect estimates, finding that mortality decline is strongly associated with falling absolute persistence but only weakly associated with falling proportional persistence. A reduction in under-5 mortality of 0.1 (100 per 1000 live births) is associated with a weakening of the average marginal effect by 0.007, large relative to the mean average marginal effect of 0.009. If we run a regression of UN under-5 mortality on period and country fixed effects in our sample, the period fixed effects indicate that mortality fell on average by 0.152 (i.e., 152 per 1000 live births) over the 40 years in our sample. Multiplied by our coefficients, that change implies a decline in the average marginal effect by 0.010. In contrast, log GDP per capita and armed conflict exhibit no relationship with mortality persistence.

At the same time, net of country and period fixed effects, the odds ratio is not significantly related to the under-5 mortality rate, although the coefficient is also positive. Aggregate mortality decline tends to benefit high-mortality lineages more than low-mortality lineages in terms of absolute bereavement risk. But in proportional terms, high-mortality lineages do not gain significantly on low-mortality lineages. In this sense, relative inequality in bereavement risk does

[13] Our series from the UN Population Division peaks at 466 per 1000 live births in 1990-4, while the UN Inter-Agency Group for Child Mortality Estimation series peaks at 276 in 1994 (http://childmortality.org).

not diminish during the process of aggregate mortality decline.

The Appendix extends Table 1.4 in two dimensions. First, Figure A.10 reports estimates of a semi-parametric version of the two-way fixed effect regression, in which the under-5 mortality rate enters as a series of bin indicators rather than a single linear term. The results show no substantial departures from linearity. Second, Figure A.11 re-estimates the regression leaving out one country at a time, showing that our estimates are not driven by any single country.

1.4 Conclusion

In data from a broad swath of the developing world, risk of losing a child is intergenerationally persistent. The fact of this persistence may not be surprising, but its magnitude is new, and it is large. Within a population at any given age, women with at least one sibling who died in childhood face 39% higher odds of losing at least one child. At the end of the childbearing period, these elevated odds translate to a 7 percentage point risk increase. This pattern appears to only partly reflect persistent inequality in socioeconomic variables. Geographic inequality plays a larger role in the perpetuation of child mortality risk across generations.

In studying child loss across generations, we contribute to a new demographic literature that shifts attention from deceased individuals to their survivors. Bereaved individuals face physical and mental health risks [Stroebe et al., 2007], and unequal distributions of mortality imply unequal distributions of bereavement. The cumulative toll of maternal bereavement is large; in some African countries, a majority of middle-aged women have experienced the death of at least one child under 5 [Smith-Greenaway and Trinitapoli, 2020, Alburez-Gutierrez et al., 2021, Smith-Greenaway et al., 2021]. The unequal burden of bereavement may be an important source of inequality in wellbeing, especially when it involves parental loss of a child [Li et al., 2003, 2005, Rogers et al., 2008].

The excess risk of child death among women who were bereaved of their siblings dissipates as child mortality falls in aggregate, suggesting a new way to think about the distribution of mortality decline. In absolute terms, high-mortality lineages benefited more from mortality decline. This result may be specific to the unprecedented, broad-based improvements in child health that many developing countries experienced in the late 20th and early 21st centuries. As the drivers of mortality decline shift from public health programs to individualized medicine, mortality decline may become less progressive.

Table 1.1. Descriptive Statistics, Women Aged 20-49

	Mean	Std. Dev.
Age	32.00	8.28
Years of education	4.88	4.73
Rural residence	0.63	0.48
Siblings ever born	5.67	2.64
Siblings deceased under 5	0.61	1.15
At least one sibling deceased under 5	0.32	0.47
Children ever born	3.36	2.66
Children deceased under 5	0.43	0.90
At least one child deceased under 5	0.26	0.44
Observations	1,288,072	

Note: Sample includes women with at least one sibling ever born from 119 Demographic and Health Surveys in 44 countries. Sampling weights are rescaled to reflect each survey's contribution to the sample.

Table 1.2. Pooled Estimates of Mortality Persistence

	Logit (Mother) Any child death		Poisson (Mother) # child deaths		Logit (Birth) Child death	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Indicator of sibling death						
Any sibling U5 death	1.386	1.370	1.237	1.242	1.202	1.233
	[.0085]	[.009]	[.0058]	[.0062]	[.0069]	[.0077]
Sibs ever born		1.005		.998		.988
		[.0012]		[.00094]		[.0012]
AME(any sib. death)	.073	.070	.219	.224	.021	.024
Observations	1,288,072	1,288,072	1,288,072	1,288,072	2,609,862	2,609,862
B. Count of sibling deaths						
# sibling U5 deaths	1.141	1.141	1.085	1.093	1.075	1.092
	[.0027]	[.003]	[.0017]	[.002]	[.0022]	[.0025]
Sibs ever born		1.000		.992		.982
		[.0012]		[.00098]		[.0012]
AME(# sib. deaths)	.029	.029	.084	.091	.008	.010
Observations	1,288,072	1,288,072	1,288,072	1,288,072	2,609,862	2,609,862

Note: The reported estimates are logit odds ratios and Poisson incidence rate ratios. Brackets contain standard errors clustered at the survey cluster level. AME refers to the average marginal effect of sibling death(s); in the woman-level models, it is computed at age 49. All models include survey indicators. Woman-level models also include indicators for the woman's age in single years. Sampling weights are rescaled to reflect each survey's contribution to the sample.

Table 1.3. Adding Covariates

	Logit with survey effects	Conditional logit with cluster effects	
	(1)	(2)	(3)
A. All, N = 2,397,677			
Sibs deceased under 5	1.085	1.043	1.037
	[.0025]	[.0024]	[.0024]
Sibs ever born	Yes	Yes	Yes
SES variables	No	No	Yes
B. Non-migrants, N = 491,640			
Sibs deceased under 5	1.073	1.041	1.038
	[.0047]	[.0051]	[.0051]
Sibs ever born	Yes	Yes	Yes
SES variables	No	No	Yes
C. Not missing height, N = 1,191,027			
Sibs deceased under 5	1.077	1.038	1.037
	[.0033]	[.0033]	[.0033]
Sibs ever born	Yes	Yes	Yes
Height	No	No	Yes

Note: This table reports odds ratios from birth-level logit regressions of under-5 child death on the number of under-5 sibling deaths, the number of siblings ever born, and the indicated explanatory variables. We omit clusters lacking variation in child mortality. Migrant status is available in 77 surveys; height is available in 103. Socioeconomic variables include maternal education and a wealth index based on principal component analysis over a vector of durable goods ownership indicators. Sampling weights are rescaled to reflect each survey's contribution to the sample.

Table 1.4. Panel Analyses of Mortality Persistence over the Mortality Transition

	Average marginal effect				Odds ratio			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Under-5 mortality [0-1] (UN)	0.065**			0.068**	0.54			0.59
	[0.024]			[0.025]	[0.36]			[0.38]
Log GDPpc (PWT)		0.00044		0.0013		0.023		0.030
		[0.0016]		[0.0014]		[0.018]		[0.018]
Conflict [0/1] (UCDP/PRIO)			-0.0000072	0.00050			-0.0036	0.0019
			[0.0014]	[0.0013]			[0.014]	[0.012]
Mean of estimated parameter	0.009	0.009	0.009	0.009	1.092	1.092	1.092	1.092
Std. dev. of estimated parameter	0.008	0.008	0.008	0.008	0.090	0.090	0.090	0.090
Country-period cells	270	270	270	270	270	270	270	270

Note: Each observation is a country-period cell. The dependent variable is the cell-specific mortality persistence odds ratio (OR) or average marginal effect (AME) estimated from a birth-level logit regression of under-5 death on the mother's number of under-5 sibling death, the mother's number of siblings ever born, and survey indicators. All cell-level regressions include country and period fixed effects. Brackets contain standard errors clustered at the country level.

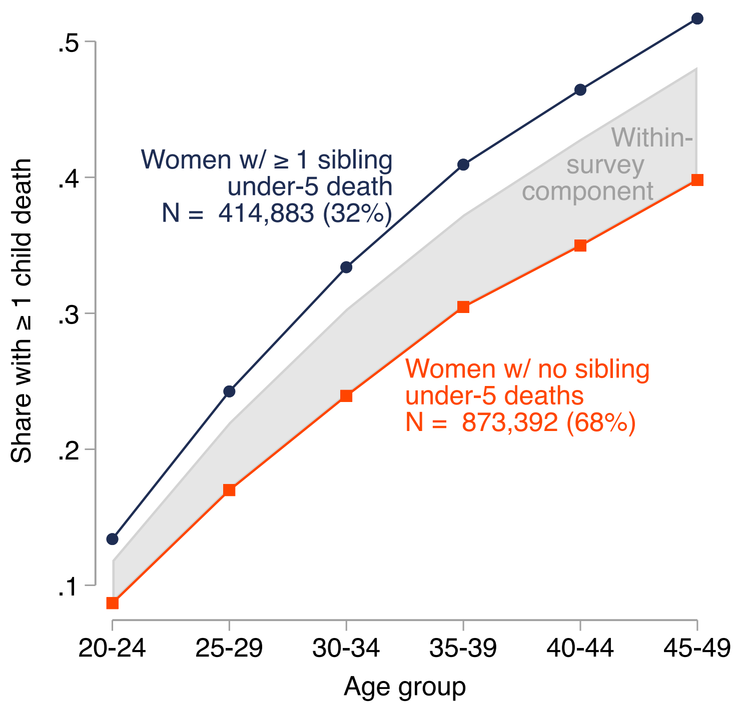


Figure 1.1. Share with Any Child Death, by Any Sibling Death
Note: For each five-year age group, we plot the share of women with at least one child death separately for women with and without deceased siblings. The within-survey component is the weighted average of the difference in shares within in each survey. Sampling weights are rescaled to reflect each survey's contribution to the sample.

Chapter 1, in full, is a reprint of the material as it appears in American Economic Review: Insights. Lu, F., & Vogl, T. (2023). Intergenerational persistence in child mortality. American Economic Review: Insights, 5(1), 93-109. The dissertation author was a primary investigator and co-author of this paper.

Chapter 2

Marriage outcomes of displaced women

2.1 Introduction

Displacement generates tremendous risks: from loss of life and property, to the possibility of long term physical and psychological harm. Empirical research on the impacts of displacement on individuals is notoriously difficult because of issues such as non-random selection of displaced populations, lack of data during periods of displacement, selection due to mortality, choice of destination, etc. Yet, such research is undoubtedly in need as millions of people are displaced every year [UNHCR, 2019], and millions are at risk of displacement due to political, economic, or climate related reasons [Steele et al., 2007, iDMC, 2017]. Women are made particularly vulnerable by displacement.[1] We study the marriage market outcomes of displaced women, as marriage-related outcomes (such as age at marriage and partner characteristics) have been shown to impact human capital investments and long term fertility outcomes [Jensen and Thornton, 2003, Field and Ambrus, 2008].

The effect of displacement on marriage outcomes is theoretically ambiguous. There are reasons to think that displacement could induce earlier marriage for displaced women [Corno et al., 2020, Mourtada et al., 2017, Schlecht et al., 2013, Nour, 2009]. For example,

[1]For example, women who are displaced, whether internationally or internally, are at a higher risk for sexual and gender based violence [UNHCR, 2003, Brookings Institution, 2014]. Reports from various countries with large portions of internally displaced women, like Syria or Iraq, suggest a lack of protection and high degree of violence towards them [Bradley, 2013].

married women may face fewer threats due to the presence of a male spouse in the household, displaced parents may seek marriage of young daughters to alleviate financial constraints, or earlier marriage of young women might be a way to quickly assimilate with local populations. At the same time, other marriage market mechanisms suggest that displacement could delay marriage for displaced women. Displacement may increase the search cost of finding good matches if families have less information or limited social networks in their new destinations [Oppenheimer, 1988]. Additionally, displaced families may not want their daughters to marry if they are productive household members in home production or the labor market. Additionally, the nature of displacement may generate asymmetries in the supply of men versus women, and if displaced populations tend to marry within their own communities, then this factor could lead to upward or downward pressure on the age of marriage of women as well. Hence, an empirical investigation into the marriage outcomes of displaced women is required in order to begin to consider which of the above mechanisms might be at play.

We begin by documenting marriage patterns among displaced versus non-displaced women using 12 representative survey datasets from 7 countries. Examining marriage rates at a given age, we find remarkably consistent results across nearly all the countries in our sample: being young and displaced leads to earlier marriage than being young and not displaced. While this result is certainly fraught with the usual concerns of who is displaced and who is not displaced, this pattern is robustly observed across a wide range of countries and years. We conduct the same analysis for displaced and non-displaced *men* and find little evidence of the same pattern; displaced and non-displaced men, regardless of age, seem to marry at similar rates.

Having established this broad pattern in the data, we turn to the specific case of the partition of India and formation of Pakistan in 1947, where the data and context allow us to answer this question while accounting for some of the data limitations and selection problems in the cross-country correlations. In this case, we not only observe when most people were displaced (i.e. 1947-49), but we also know their precise age at marriage. Moreover, the issue of selection with regards to who gets displaced and where they end up is mitigated (although not fully solved), since our data examines refugees in Pakistani Punjab and nearly *all* Muslims from Indian Punjab were forcibly displaced around the time of partition (the fraction Muslim in Indian Punjab went from 30% in 1931 to 1.75% by 1951) and nearly all of them settled close to the border in Pakistani Punjab and in districts vacated by displaced Hindus [Bharadwaj et al., 2008]. This allows us to compare the age at which young displaced women married relative not just to young non-displaced women at their destinations, but also relative to displaced women who were older and hence already married before partition. Comparing women from the same group who are older at the time of displacement introduces a plausible counterfactual to those who are young at the time of displacement. In that sense, our empirical analysis is a standard difference in differences design: we compare marriage outcomes of women who are young at the time of displacement to women who are older at the time of displacement and to women who are never displaced.[2] There are some potential concerns with using this strategy such as selective mortality, the role of out-migrating Hindus, marriage market effects on non-displaced women, and measurement of migration *after* partition. We discuss these in detail in Section 2.5.1 of the paper.

The results from Pakistan paint a similar picture to the overall cross-country evidence. We focus on the province of Punjab that received a bulk of partition-related migrants. Displaced women who were adolescents (between the ages of 13 and 17) at the time of partition married 0.28 years earlier (average age of marriage for the control group is around 19 years), were 3.8 percentage points more likely to marry before the age of 18 (control mean is 30.8%), and are 5.3 percentage points more likely to be married during partition (defined as marrying between years 1947 - 1949; comparison mean is 1.7%). To what extent are these results specific to *adolescent* women during partition, or do they simply reflect a tendency for displaced women, regardless of age, to get married earlier? To get a sense of this, we examine outcomes for displaced women who were very young children (between the ages of 1 and 5) at the time of partition. For this group, we find small negative and statistically insignificant effects on age

[2]Due to data limitations, we are unable to examine marriage outcomes for displaced men in this setting.

of marriage and likelihood of marriage before 18. Taken together, the cross-country evidence and the more specific data from the case study of the partition suggest that displaced women tend to marry earlier than non-displaced women, but that these impacts depend on the timing of displacement during the lifecycle.

While we are not able to disentangle the impact of displacement from that of early marriage in this context, examination of additional outcomes reveals that displaced women between the age of 13 and 17 at the time of partition were less likely to continue their education and had higher fertility (statistically significant effect of 0.43 more children born over a control mean of 5.6). On the other hand, displaced women who were much younger (between the ages of 1 and 5) at partition were 3.5 percentage points more likely to be literate (control mean of 2.8%), 3.1 percentage points more likely to have completed primary school (control mean of 2.1%), and have more surviving children. The data also allows us to examine spousal quality. For adolescent women, we find no impacts on the characteristics of who they marry, but young children at the time of partition appear to have married more educated and younger men. The results for both age groups are broadly consistent with the existing literature on the relationship between marriage timing, fertility and education attainment [Jensen and Thornton, 2003, Field and Ambrus, 2008].

Our paper contributes to a robust literature examining how displacement affects women, although most of it is outside the field of economics. Many papers in this area focus on the violence that women face during periods of displacement [Vu et al., 2014, Usta et al., 2008, Masterson et al., 2014, Keygnaert et al., 2012, Wirtz et al., 2014], the challenges women face in accessing healthcare [Gagnon et al., 2002, Usta and Masterson, 2015, Morris et al., 2009], and the disruption of education and learning that has significantly larger impacts on girls [Sirin and Rogers-Sirin, 2015]. These papers in turn relate to a broader literature on refugees and the unique experiences and challenges of female refugees (e.g. [Becker et al., 2020, Sieverding et al., 2018, Fasani et al., 2018, Maia et al., 2018]).[3] Our paper also relates to and is broadly consistent with

[3] See Becker and Ferrara [2019] for a broader review on the consequences of forced migration.

work examining the links between early marriage and later life outcomes for women [Jensen and Thornton, 2003, Field and Ambrus, 2008, Vogl, 2013, Chari et al., 2017, de Groot et al., 2018, Raj et al., 2019]. In examining the particular case of the partition of India, we add to a recent empirical literature examining the effects of the partition on demographic and other outcomes such as conflict, trade, and agriculture [Bharadwaj et al., 2008, Jha and Wilkinson, 2012, Bharadwaj and Fenske, 2012, Bharadwaj and Mirza, 2019]. Additionally, we contribute empirical evidence to a large literature on the impact of the partition of India on women [Bhasin and Menon, 1998, Butalia, 2017, Das, 2006].

The remainder of the paper is organized as follows: Section 2.2 provides data, empirical strategy and results for the cross-country evidence on marriage outcomes of displaced women, and sections 3 onwards delve into this question for the specific case of the partition. Section 2.3 provides background and context; sections 3.2 and 2.5 describe the data and empirical strategy respectively; section 2.6 reports results; section 2.7 provide a discussion and conclude the paper.

2.2 Cross-country Evidence

2.2.1 Data

We use cross-sectional representative survey data from various countries (microdata samples are publicly available through IPUMS International[4]) that allows us to identify displaced persons. These countries (with survey rounds in the listed years) are: Armenia (2011), Cambodia (1998, 2004, 2008, 2013), Colombia (2005), India (1983, 1987, 1999), Iraq (1997), Kyrgyz Republic (2009), and Nepal (2011). For all surveys, the respondent's marital status (never married, currently married, separated/divorced, widowed) was measured at the time of enumeration. Most surveys don't include the respondent's age at first marriage, with the exception of Cambodia (2004, 2013), Iraq (1997), and Nepal (2011). For these surveys, the respondent's age at first marriage was asked retrospectively at the time of enumeration.

[4]Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D020.V7.2

The universe of migrants identified is different across surveys. All datasets, except those from Nepal (2011) and Colombia (2005), identify migrants as those who have ever lived outside of their locality (village, town, city) of enumeration.[5] The Nepal (2011) dataset identifies migrants as those who have ever lived outside of their district of enumeration. The Colombia (2005) dataset identifies migrants as those who have lived outside their locality of enumeration in the past 5 years. Due to lack of information on the locality of birth,[6] we cannot identify respondents who have ever lived outside of their locality of enumeration. Therefore, for the Colombia (2005) dataset we are constrained to use the set of migrants that are identified by the survey data, which means we exclude any migrants who migrated more than 5 years ago from the migrant sample and include them in the non-migrant sample.

All surveys document the cause of the respondent's most recent migration. We use this information to define displaced people as the subset of migrants whose cause of migration was insecurity, war, violence, or natural disaster. The surveys also ask for the number of years living in the locality of enumeration, which we interpret as years since the most recent migration for the migrant sample. We use this to define recently displaced as a subset of displaced who migrated up to 5 years before the time of the survey. Finally, we define non-migrants as those who have never migrated according to their survey response.

Table 2.1 reports summary statistics on sample composition and marriage outcomes by country. Across all datasets used in the cross-country analysis, the proportion of the population that has last migrated due to displacement is very similar for males and females, and accounts for a small proportion of each country's total population (between 0.1 and 1.5 percent). The proportion of non-migrants differs by gender in some countries, notably in contexts with higher levels of female marriage migration (India and Nepal).

Across countries that measure marriage age, we don't see a consistent pattern between

[5]For all India datasets, the survey codebooks specify that the person must have lived in their last location for at least 6 months.

[6]The data contains municipality of birth, but we cannot use this to identify non-migrants, as most displaced persons are still living in their municipality of birth.

the average age of marriage of displaced and non-migrant women. However, this doesn't take into account the timing of displacement for displaced women or the age distributions of each group. When narrowing our attention to recently displaced women, we see that young recently displaced women generally have higher marriage rates than young non-migrant women.

2.2.2 Empirical strategy

Using the 12 representative datasets across 7 different countries, we document differences in age-specific marriage rates between recently displaced women and non-migrant women comparison groups as an informative descriptive exercise.[7]

There are several challenges in working with these population surveys. First, within all datasets, displacement occurs over long periods of time, which leads to concerns regarding the endogeneity of the timing of displacement. Second, for migrants identified by each dataset, we only have information on the reason for their most recent migration. Therefore, we cannot identify displaced people who previously migrated for other reasons (e.g. marriage) or people who have been displaced multiple times. Third, in a majority of our datasets, we don't know the respondent's age at first marriage, which makes it difficult to measure the impact of displacement on marriage. For people who were displaced long before the time they were surveyed, information on whether they ever married at the time of enumeration does not allow us to identify whether displacement impacted their marriage timing. Finally, there are likely to be many observable and unobservable characteristics that lead to non-random selection of displaced populations. Therefore, we do not claim that observed patterns in the differences in age-specific marriage rates in this analysis are solely due to the impact of displacement.

We study recently displaced individuals (those who have been displaced in the 5 years before their time of the survey) to circumvent the issue that we can only observe the reason for an individual's most recent migration. Furthermore, we expect the observed impact of displacement

[7] We exclude the 1973 Pakistan dataset from this exercise for several reasons. The partition of India occurred 25 to 26 years prior to survey, so the displacement event is not recent enough to be used for the purpose of this exercise.

on marriage to dissipate as the time since displacement (and mechanically the age of the displaced women) increases. However, this approach does not address the other identification concerns listed above.

The dependent variable we consider is whether the respondent was ever married (those who are currently married, separated/divorced, widowed). We calculate marriage rates for recently displaced and non-migrant women by age at the time of the survey. We choose non-migrant individuals for our comparison group, as by definition they could not have been displaced in the past[8].

To examine variation in the relationship between recent displacement and marriage formation by age, we estimate

$$Married_{iadct} = \alpha_{ct} + \sum_{a \in A} \delta_a Age_a * Displaced_d + \sum_{a \in A} \beta_a Age_a + \gamma Displaced_d + \varepsilon_{iadct} \tag{2.1}$$

where $Married_{iadct}$ is an indicator of whether individual i with age at survey a and displacement status d in country c with year of survey t was ever married. α_{ct} is a country by year fixed effect. Age_a is an indicator of whether the individual is age a at time of the survey. $Displaced_d$ is an indicator of whether the individual was recently displaced (displaced within 5 years of time of the survey). The omitted category for "age at survey" is 38 year olds. The sample is restricted to recently displaced and non-migrant women (or men) with age at survey between 10 and 40. Results from this specification are reported in Figure 2.1 using data pooled over all countries. To examine the robustness of the estimates from Equation 2.1, we replace the country-by-year fixed effect with more granular destination by country-by-year and birthplace-by-country-by-year fixed effects. Additionally, we report all results with and without clustered standard errors.

[8] One potential concern with this comparison group is that it excludes individuals who migrated due to marriage, which is very common for women in countries such as India, so non-migrant individuals may mechanically have lower marriage rates. This concern also applies to our defined group of recently displaced women, since our sample definition will exclude people who migrated due to marriage after displacement.

2.2.3 Results

Figure 2.1a reports estimates for the difference in marriage rates between recently displaced and non-migrant people estimated by Equation 2.2 by gender using pooled data from all countries except Pakistan. We find that women who are displaced at young ages show significantly higher marriage rates than non-migrant women of the same age, indicating that displaced women get married at younger ages. Marriage rates for recently displaced and non-migrants converge to similar levels in older age groups (can also be seen in panel (b) of Table 2.1). We conduct the same analysis for men but do not find evidence of the same pattern (Figure 2.1b). Displaced and non-migrant men are married at similar rates for all ages.

Results for both men and women are robust across specifications with more granular destination by country by year (Appendix Figure B.1) and birthplace by country by year fixed effects.[9] All cross-country results are also robust to clustering standard errors at the geography by year by displaced level, where geography by year is the level of the fixed effects in the given specification. This pattern is also similar across individual countries.

Having shown strong correlations that demonstrate higher marriage rates among young displaced women across different countries and time periods, we now strengthen the empirical narrative by considering the specific case study of displacement into Pakistan caused by the partition of India in 1947.

Partition of India in 1947

2.3 Background

The end of British rule in India in August 1947 led to the partition of India into two countries: India (Hindu-majority state) and Pakistan (Muslim-majority state). The partition led to one of the largest and most rapid migrations in human history. It is estimated that about 14.5

[9]To include birthplace by country by year fixed effects, we are restricted to using datasets that include nativity information and displaced individuals who were displaced within their country of origin. All datasets in the cross-country sample include nativity information, with the exception of the India datasets (1983, 1987, 1999).

million people were displaced along religious lines by 1951, with 6.5 million displaced into Pakistan [Bharadwaj et al., 2008].

The eastern provinces of Pakistan, Punjab (to the north-east) and Sindh (to the south-east), were the epicenter of migration activity in the country which saw an estimated outflow of 8 million Hindus and Sikhs and an estimated inflow of nearly 6.5 million Muslims by 1951 [Bharadwaj et al., 2008]. With respect to Punjab, high levels of ethnic and religious violence caused forced expulsion of minorities from either side of the border. The Governor of West (Pakistani) Punjab at the time of partition estimated that 500,000 Muslims died trying to enter his province [James, 2010]. The violence and forced displacement meant that at least in some areas migration was unlikely to be selective: between 1931 and 1951, the percentage of Muslims in East (Indian) Punjab fell from 30% to 1.75% and the percentage of Hindus/Sikhs in West Punjab fell from 21.7% to 0.16% [Bharadwaj et al., 2008]. Nearly all Muslim migrants from East (Indian) Punjab migrated to the contiguous West (Pakistani) Punjab, settling in an ethnically, culturally and linguistically similar environment. These migrants accounted for the vast majority (more than 80%) of migrants in West Punjab [Chitkara, 1998]; in fact, according to the 1951 Census of Pakistan, Muslim refugees from East Punjab and nearby princely states constituted 80.1% of Pakistan's total refugee population [Chitkara, 1998] in 1951.

We focus on Muslim women who were displaced into the Punjab province of Pakistan due to the attributes of migrants and the migration into (West) Punjab relative to those in Sindh.[10] The sudden inflow of non-selected migrants who were culturally, ethnically and linguistically similar people from Indian Punjab as a result of partition means that West (Pakistani) Punjab provides a unique opportunity to strengthen the evidence on the impact of displacement on

[10]Refugee movement into Sindh province was very different. Prior to partition, Karachi, a port city and the largest urban area in Sindh, was host to an educated Hindu population and a less educated and less prosperous Muslim population. After partition, Karachi became the capital of Pakistan and attracted Muslim migrants from central and southern states of India such as Uttar Pradesh, Bhopal, Bombay, Gujarat, Madras and Mysore to replace the majority out-migrating Hindu population. These migrants were more educated, and ethnically, culturally and linguistically different from the locals [Tan et al., 2000]. Overall, Sindh received nearly 1.1 million migrants by 1951, half of whom travelled to Karachi [Jaffrelot, 2015] followed by a steady inflow of migrants in the 1950s and 60s also from the central and southern states of India [Khalidi, 1998].

marriage outcomes found in the the cross-country data.

Using data from Pakistan's 1973 Housing, Economic, Demographic Characteristics (HED) survey (detailed further in Section 3.2) we plot the year of migration for Indian-born women in Punjab in Figure 2.2, which shows a large spike in migration at the time of partition (years 1947 - 1949). For those who migrated after partition, we cannot distinguish between those who are re-migrating within Pakistan or migrating from India to Pakistan for the first time. Historically, we know there was a continued flow of Indian-born Muslim migrants to Pakistan after partition, however the majority of migrants went to Karachi and other cities in Sindh [Khalidi, 1998].

Finally, displaced women have a different spatial distribution from native-born Pakistani women in Punjab. As seen in Figure 2.3, they are more likely to live near the Indian border and urban areas (Faisalabad and Lahore). 37.7% of displaced women live in urban areas, while 18.3% of native-born women live in urban areas. Bharadwaj et al. [2008] documents that district-level population inflow into Pakistani districts was highly correlated with outflows and had an approximately 1-1 relationship.

2.4 Data

The data we use are from Pakistan's 1973 Housing, Economic, Demographic Characteristics (HED) survey (microdata sample is publicly available through IPUMS International). Approximately 24,000 blocks (sub-districts) were sampled out of 75,000 in the country. A sample of households was taken from each sampled block to yield a sample size of 300,000 households.[11] Urban households were oversampled relative to rural households. Roughly 15% of households in the dataset do not have a head and appear to be fragmented. As our study focuses only on the context of Punjab (from this point onwards, unless explicitly stated, Punjab refers to Pakistani (West) Punjab), we subset this data to a subsample of nearly 193,000 households in Punjab.

[11] According to data documentation, there are 300,000 households, but our data has 322,131 unique households.

The survey collected data on the demographic characteristics of individual household members, including marital status and migration history. The respondent's marital status (never married, currently married, separated/divorced, widowed) was measured at the time of enumeration. For females, the respondent's age at first marriage was asked retrospectively at time of enumeration.[12]

We can identify a respondent's country of birth and years living in their current locality (e.g. village, town, or city) from the survey. We use this information to classify individuals as those displaced by partition, native-born, and non-migrants. We define individuals displaced by partition as those who were alive at the time of partition (age 25 or older at the time of the survey) were born in India and have been living in their current locality in Pakistan for 24 to 26 years before enumeration (migrated to Pakistan between the years of 1947 and 1949).[13] We are not able to identify displaced individuals who subsequently migrated within Pakistan. We define native-born as being born in Pakistan (in Punjab, 88.9% of native-born women who were alive at partition are born in Punjab), and non-migrants as those who reported that their years of living in their current locality is greater than or equal to their age at the time of enumeration. By definition, non-migrant women are a subset of native-born women, so these categories are not mutually exclusive.

Additionally, we observe education and fertility outcomes. For education, we observe whether the respondent is literate and their highest level of education attained. The levels of education recorded in that data are no education, some primary, completed primary (grade 5), completed middle (grade 8), completed matriculation (grade 10), completed higher secondary education (grade 12), and other levels for higher education. For each level of education attainment, we calculate the minimum age needed to obtain that level of education (5 plus years of schooling). If the minimum age needed to obtain that schooling level is higher than age at partition, we

[12]In the entire dataset, this variable is missing for 97.1% of ever married men, but for only 1.6% of ever married women. Therefore, we are unable to conduct analysis for men using this variable. The enumeration codebook only details how to ask this question to ever married female respondents.

[13]Our results are robust to including women who migrated in later years, up to the year of 1952.

classify the respondent as having continued schooling after partition. For fertility, we observe the respondent's number of children ever born and surviving at time of enumeration.

The last two rows of Table 2.1 report summary statistics on sample composition and marriage outcomes for Pakistan and Punjab respectively. In contrast to other countries, people displaced from India at time of partition constituted a substantial 5.2 (5.8) percent of the Pakistan female (male) population and 6.5 (7.4) percent of the Punjab female (male) population in 1973.[14] Nearly 30% of individuals aged 26 or older at the time of the survey (that is, who were born at or before the time of partition) are displaced.

2.5 Empirical Strategy

Our identification strategy is a difference in differences across displacement status and age at partition. We compare young displaced women to two groups: older displaced women and native-born women. Older displaced women whose demographic characteristics (marriage, education, fertility) were largely fixed at the time of partition serve as a comparison group to younger displaced women. Native-born women provide counterfactual age-cohort demographic trends.

We thus estimate:

$$Y_{iad} = \alpha + \delta Young_a * Displaced_d + \beta Young_a + \gamma Displaced_d + \varepsilon_{iad} \tag{2.2}$$

where Y_{iad} is an outcome for an individual i in age group a at partition and displacement status d. $Young_a$ is an indicator of whether the female was young (between age 13 to 17) at the time of partition. $Displaced_d$ is an indicator of whether the female was displaced from India at the time of partition (as defined in Section 3.2).

The main analysis sample includes all displaced and native-born (comparison) group

[14] The fact that everyone born in Pakistan since 1947 is, by definition, a native-born lowers the proportion of partition migrants in the total population.

women aged 13 to 17 and 30 to 32 at the time of partition (38 to 42 and 55 to 57 at time of enumeration) and living in the state of Punjab at the time of enumeration.

We chose women who were 13 to 17 years old at partition as our primary "young" age group for several reasons. The upper limit of the age of 17 allows us to analyze the impact of displacement on marriage under the age of 18 (the most common definition of "child marriage"). We choose the age of 13 as the lower limit because this is an age by which many women have reached puberty.[15] Additionally, defining the lower bound as 13 year olds allows us to cleanly classify educational attainment before and after partition. It is very unlikely that women who were 13 years or older at partition had their *primary* schooling interrupted by partition, so we can attribute any primary education completion for this group as occurring prior to partition. To better understand how the timing of displacement impacts the outcomes of displaced women, we also study impacts on women who were young at the time of partition. For this analysis, we define the "young" group as women who were between the ages 1 and 5 at partition. This group also allows us to cleanly identify the effects of displacement on education accumulation, as all educational attainment for this age group must have occurred after partition. Our results are robust to different definitions of both levels of the "young" age group. Results are nearly identical if we classify the two "young" analysis groups as 10 to 17 (or 10 to 19) and 1 to 9 years old at partition. For completeness, we report results with these alternative definitions and for women who were ages 6 to 12 at the time of partition in Appendix Table B.4.

There are two opposing forces that we were concerned about when choosing the "old" comparison group: the partial treatment on marriage outcomes for younger age groups and (potential) selective mortality for older ages. We choose women aged 30 to 32 at partition as our older comparison group because their characteristics (marriage age, fertility and education) were much less likely to be impacted by partition.[16] We choose to exclude women over the age of 32

[15] Puberty can be formally thought of as age at first menarche, which is an important determinant of marriage timing [Field and Ambrus, 2008].

[16] Women under the age of 30 at the time of partition are "partially treated" to a much larger extent than women ages 30 or above. At ages of 25 and 26, the proportion of women married before partition is less than 90%. The proportion married jumps from ~92-93% at age 29 to ~96-97% at age 30. Above the age of 30, the proportion

at partition due to concerns of potential differential mortality selection at older ages although our main results are extremely robust to including women up to the age of 50 at the time of partition in the older comparison group (results reported in Appendix Table B.5).

We report the average value of each outcome variable for "old" native-born women as the comparison mean. While we use non-migrant women as our control group in the cross-country analysis, we prefer to use native-born women as the comparison group in this context to include (Pakistan-born) women who may have migrated for marriage or other reasons within Pakistan. Our results are highly robust to the choice of control group and do not change substantially when we use non-migrant women as a control group instead. Our identification strategy relies on the assumption that displaced and native-born women have the same counterfactual age-cohort trends in outcomes in the absence of partition.

One might expect displaced women's outcomes to vary by their destinations, particularly for women displaced as children. As a robustness check, we control for destination (district by urban) fixed effects in Equation 2.2. Our results are very similar under this specification.[17]

To conduct inference, we use heteroskedasticity-robust standard errors in our preferred specification. In addition, we report results using standard errors clustered at the interaction of age group (young or old at partition), whether displaced (displaced or native-born), district of residence (19 districts in Punjab), and urban (whether lives in an urban or rural area) level in Appendix Tables B.1, B.2, and B.3. There are three assumptions required for this clustering approach to be justified. First, unobserved shocks to marriage markets between destinations (defined at the granularity of urban or rural areas within a district) are not correlated. To the extent that marriage markets are local, this assumption is plausible. During this period in Pakistan, women were generally married to men from nearby areas [Shah et al., 1983]. Makino [2019] found that 39% of wives had the same birth village as their husbands in rural Pakistan. Second,

is typically 96% or higher. Therefore, we prefer to leave women under the age of 30 out of the main comparison group.

[17]Additionally, we conduct parallel trends tests estimated with Equation 2.3 including destination fixed effects. We do not find strong evidence of differential trends between displaced and native-born women living in the same destination at the time of the survey.

the unobserved shocks to women of different age groups within destinations are not correlated. The older women, who are almost all married by the time of partition, experienced shocks to their marriage outcomes long before the young women enter the marriage market. Finally, displaced and native-born women of the same age group experience uncorrelated unobserved shocks within destinations. This assumption is the most difficult to justify, but there is some evidence that it is plausible. To a large extent, we find that displaced and native-born women are largely marrying men from their respective groups which indicates "segregated" marriage markets. To the extent to which the communities participating these markets receive independent shocks (e.g. due to social distance or spatial segregation), this assumption may hold.[18]

2.5.1 Threats to identification

Parallel trends: We test for the parallel trends assumption by estimating a difference-in-difference specification using displaced and native-born women aged 30 to 50 at partition as the analysis sample. We estimate the following equation

$$Y_{iad} = \sum_{a \in A} \delta_a Age_a * Displaced_d + \sum_{a \in A} \beta_a Age_a + \gamma Displaced_d + \varepsilon_{iad} \tag{2.3}$$

The base group is those who were aged 30 to 32 at the time of partition. Age_a is an indicator of whether the female is in age group a at the time of partition. $Displaced_d$ is defined as above. The estimates from this specification (reported in Figure 2.6) suggest that the parallel

[18]Clustering at a more aggregated level would yield too few clusters to allow for correct inference given our empirical strategy and data limitations. Multi-way clustering is also not possible as cluster-robust standard errors are biased for a small number of clusters. The most common method to estimate cluster-robust standard errors with a small number of clusters is the wild bootstrap method, as proposed by Cameron et al. [2008]. This method has been commonly applied in difference in differenced analysis, but Canay et al. [2021] find that the additional assumptions needed for the asymptotic consistency of the wild bootstrap error estimates are difficult to satisfy for difference-in-difference studies. For our study design, the assumptions require that the average value of the independent variable of interest, in this case the interaction of young at partition and displaced, be the same across all clusters. This precludes us from estimating clustered standard errors using the wild bootstrap method that uses displacement or age at partition as dimensions for clustering. Additionally, we cannot use geographical variables for clustering using this method, as displaced women do not have a uniform spatial distribution across Punjab (Figure 2.3).

trends assumption is satisfied for our main outcomes of interest.[19]

Selective mortality: Due to the violence targeted towards women, particularly unmarried women, during partition, one may worry that mortality selection may impact the results. If partition has a significantly higher mortality toll on unmarried displaced women (and in turn the displaced young relative to the displaced old), then results showing earlier marriage due to displacement could be driven by *missing* unmarried young displaced women. To explore this, we plot the distribution of birth year for displaced and native born women who are alive at partition by whether they were married before partition (pre-1947) or unmarried at partition. If there were *missing* young unmarried displaced women, then the distribution of displaced and native-born women married before partition would look very different from the distribution of displaced and native-born women who are unmarried at partition. Figure 2.5 shows that among women married before partition and women not married at partition, the birth year distributions of displaced and native-born women look very similar. To perform a statistical test of the graphical evidence in Figure 2.5, we calculate the proportion of women married before partition for all women ages 1 to 32 at the time of partition separately for displaced and native-born women. We then run the following regression, where there are 64 observations (one for each age and displacement status group):

$$ProportionMarriedBeforePartition_{ad} = \alpha_a + \beta Displaced_d + \varepsilon_{ad}$$

This regression tests whether surviving displaced women have higher rates of marriage before partition than surviving native-born women. The β estimate is 0.0077 and the p-value is 0.125. This estimate is not statistically nor economically significant. This provides some evidence that mortality is not correlated with marriage timing and increases our confidence that

[19]The adherence to parallel trends in some of the older age categories (ages 37 to 39 and 47 to 49) for education outcomes is less robust than that for marriage and fertility outcomes. This may be due to differential sorting by education into these age categories due to imperfect recall, particularly for older and less educated women. This is suggested by the relatively low number of observations in these bins and the (subsequently) larger standard error estimates. However, we are not particularly concerned about these deviations for two reasons. First, if these deviations from parallel trends are due to differences in reporting of age, we do not expect this issue to be as substantial for younger women. Second, the deviations from parallel trends in the older age group are not systematic across age groups and are in the opposite direction of the effects we observe in our main age groups of interest.

selective mortality at the time of partition is not driving our results.

Marriage market effects on non-displaced women: It is also plausible that the presence of displaced women and the occurrence of partition impacted the outcomes of native-born Punjabi Pakistani women. Therefore, the effect we estimate will be net of any spillovers on the host population (in particular, spillovers on young native-born women); however, as we discuss in the results section, it appears that displaced adolescent women tended to marry displaced men and vice versa, mitigating concerns that displacement changed the marriage market via competition for native born men. Moreover, to the extent that native born adolescent women faced less economic hardship relative to displaced adolescent women, any effects of economic hardship on marriage rates of native born women would likely lead to a lower bound effect (as in Corno et al. [2020]) .

Post-Partition migration: As mentioned above, we use self-reported *years residing in current locality* in the 1973 HED survey to calculate year of migration for Indian-born women, which we then use to define displacement. Migrants between 1947 and 1949 formed a vast majority of partition-related migrants; however, while a small fraction of the overall flows, there were still migrants from India to Pakistan in the 1950s and 60s. Unfortunately we are not able to ascertain exactly how much of this post-1949 migration is associated with partition (and thus originating from India) relative to re-migration of displaced women within Pakistan (see Figure 2.2).

We choose to exclude the post-1949 Indian-born migrants from our definition of displaced (and the analysis) because we know from historical accounts these migrants most likely originate from India, rather than re-migrating within Pakistan; even within India they largely originate from central and southern states of India (as opposed to Indian Punjab) [Khalidi, 1998]; they are a self-selected group from among the remaining Muslims in those states, and their migration timing is very likely endogenous. That said, our results do not hinge on the choice of 1949 as the cutoff point. Including Indian born women who report their last year of migration up to 1953 does not meaningfully alter our results.

Outmigration of Hindus: The partition was an episode of population transfer. As Muslims from Indian Punjab came in, Hindus from Pakistani Punjab were forced to leave. However, the effects of Hindus leaving is unlikely to directly impact marriage market outcomes for Muslim women (displaced or not) since marriage rates across religious lines are extremely low. Using data from India in 2005, Goli et al. [2013] estimate only 0.6% of Muslim women married non-Muslims. Hence, it is safe to assume that in the 1940's and 1950's, rates of inter-religious marriage were also substantially low.

Destination selection: Since nearly all Muslims from Indian Punjab left for Pakistani Punjab after partition and nearly all Hindus left Pakistani Punjab for Indian Punjab, "selective" in and out migration is less of an issue in this context. However, one might still be concerned that where displaced people end up is selective, and that selection on the basis of place drives the results rather than displacement itself. We address this in two ways. First, we rely on the evidence in Bharadwaj et al. [2008] who point out that most of the refugee movement into Pakistan was determined by distance from the border and places where Hindus out-migrated from. This is unlikely to result in selection of place that somehow directly affects marriage market outcomes for adolescent displaced women. Second, as a robustness check, we control for destination (district by urban) fixed effects in Equation 2.2. Our results are almost identical under this specification mitigating concerns about destination selection. However, since destination can still be considered a "choice" or an outcome in this instance, our baseline specification does not include these fixed effects.

2.6 Results

Marriage outcomes

We first check whether women in each of the relevant groups for analysis (defined by the interaction of young (aged 13 to 17) at partition and displaced by partition) differ in their extensive margins of whether they have ever been married at the time of the survey (25 years after partition). Overall, 98.9% of women in the main analysis sample have ever been married by the time of the survey. Young native-born women have a statistically significant lower marriage rate than other groups, with a marriage rate that is 1.3 percentage points lower than that of older native-born women (significant at the 1% level) (see Table 2.2a). Even so, all groups have numerically similar marriage rates by the time of the survey at about 99%, which indicates that we should not be very concerned by this extensive margin issue. In particular, our primary variable of interest, age at first marriage, should be defined for almost all women in the sample. This is confirmed in the data. We observe the age at first marriage for 98.6% of ever married women in the main analysis sample.

Adolescent displaced women who were between the age of 13 and 17 at the time of partition (defined as "young" in our main analysis sample) have significantly higher child marriage (defined as married before the age of 18) rates and lower marriage ages. Table 2.2a reports the effect of displacement on marriage outcomes estimated by Equation 2.2. Young displaced females experienced an 3.8 percentage point (11.4% of comparison mean) increase in child marriage rate. Displacement at a young age also led to a 0.28 year (comparison mean of 19.3) decrease in marriage age.

In particular, adolescent displaced women were 5.3 percentage points more likely to get married at the time of partition (defined as being married for the first time between years 1947 to 1949). We observe an excess density mass of women who were married at the time of the event (Figure 2.4). In contrast, we observe smaller impacts on the marriage timing for women who were young children (ages 1 to 5 years old) at the time of partition (Table 2.2b).[20]

Our findings are largely robust to clustering standard errors at the interaction of the indicator of young at partition, whether displaced, district of residence, and whether the respondent

[20]At the time of the survey, women who were aged 1 to 5 years at the time of partition were 26 to 31 at the time of the survey. Women in this young group had not all been married by the time of the survey, as indicated by the lower ever married rates of young women as compared to older women. However, our results still indicate that young displaced women were getting married later. They were 1.4 percentage points less likely to be married at the time of the survey and, conditional on being married, had slightly higher marriage ages (though this result is not statistically significant).

lives in an urban area (Table B.1). The results for the timing of marriage right at partition are still statistically significant for both groups, while the effect on child marriage for adolescent displaced women is no longer significant for adolescent women (p-value = 0.388). Altogether, these results indicate that the effect of displacement on marriage outcomes in this context is driven by increased marriage at the time of displacement when adolescent girls were the most vulnerable.

At the same time, displaced women don't appear to have lower spouse quality than native-born women (Table 2.4). For adolescent women, whose marriage timing was impacted, we find no effect on spousal education or the woman's age difference with her spouse. Spouses of women who were young children at the time of partition have a smaller age difference and higher education levels. This higher spousal education level may be driven by the fact that most displaced women are marrying displaced men. Most displaced women are married to Indian-born men and vice-versa. Column 2 of Tables 2.4a and 2.4b reports effects on whether the spouse has the same nativity as the respondent. We see that both young displaced and native-born women have similar low levels of marriage to men of the opposite nativity. For women who were children at the time of partition, marriage with men of the opposite nativity increases, with an even larger change for displaced women (1.1 percentage point effect compared to 4.8 percentage point age-trend for native-born women). These results are generally robust to clustering standard errors, with the exception of the impact on spousal nativity for women who were 1 to 5 years old at the time of partition (Table B.3).

2.6.1 Other outcomes

To better understand the impact of displacement on young women, we examine their education and fertility outcomes. Literacy and primary school completion rates are higher for women who were displaced as adolescents. While we observe positive and significant effects on their literacy rate, these are almost certainly due to pre-existing differences in the education of adolescent displaced women prior to partition, as women in this age group are extremely unlikely to have completed primary school in Pakistan. While this may suggest some positive selection into who is displaced, we think this creates a lower bound effect on marriage since in general more educated women marry at older ages, relative to less educated women. In fact, adolescent displaced women were *less* likely to continue their education after partition (Column 3 of Table 2.3a). At the same time, we observe large positive effects on the educational attainment of women who were young children (1 to 5 years old) at the time of partition. As these women would not have entered primary school prior to partition, we can treat their education attainment as occurring after partition. These women experienced a 3.5 percentage point increase in literacy (comparison mean of 2.8%) and a 3.1 percentage point increase in primary school completion rates (comparison mean of 2.05%). Together, these results indicate that changes in education investments due to displacement for young females depend crucially on the timing of displacement during their lifetime. While women who were adolescents at the time of partition were less likely to continue their education after partition, we observe large increases in human capital accumulation for young children.[21]

Displaced adolescent women also have higher fertility as measured by both total children ever born and children surviving at the time of the survey. For adolescent women, displacement increased the number of children born by 0.43 (comparison mean of 5.6) and children surviving by 0.55 (comparison mean of 4.8) at the time of the survey. Higher fertility is consistent with early marriages, though we cannot disentangle the effect of displacement and early marriage in our study design. In contrast, women who were displaced as children do not bear a significantly higher number of children ever born but have a larger number of surviving children. These results are consistent with their higher education levels.[22]

[21] The positive impact on young children is similar to recent evidence from Becker et al. [2020] who, in the context of the redrawing of Poland's boundaries following World War II, find that descendants of displaced people invest more in human capital. As in Becker et al. [2020], the human capital accumulation for young children could be impacted by their parent's displacement.

[22] Our estimates may understate the impact of displacement on fertility for two reasons. First, women who were young at partition may not have completed fertility by the time of the survey. Women who were ages 1 to 17 at partition were around 26 to 43 at time of the survey. At the very least, our results show that displaced women in the 1 to 5 age group did not have children earlier than their native-born counterparts. Second, displacement may have induced older women who were ages 30 to 32 at the time of partition to have more children (for example, Nobles

2.6.2 Comparison with cross-country results

While a direct comparison of the results from the partition to that from the cross-country data sets is difficult, the two pieces of evidence offer a consistent narrative in terms of the direction of the result: displaced women tend to marry earlier than non-displaced women. The cross-country data contains different information (in the cross-country data, we do not observe age at marriage, which is a key outcome we examine in the case of the partition), but we can make still make some informed comparisons. The cross-country evidence suggests that displaced women in the 13 to 17 age range are about 12 percentage points more likely to be married than non-displaced women in the same age range. The evidence from the partition suggests that women between these ages are about 5.3 percentage points more likely to be married during the partition. While we hesitate to compare the magnitude of the effect sizes between these two exercises, the smaller effect sizes in the case of partition could be attributed to the fact that we are better able to control for underlying differences in marriage propensity (e.g. due to selection or other unobserved factors) in that context.

2.7 Discussion and Conclusion

Before we conclude, we turn to a discussion of potential mechanisms that are consistent with our results. The first mechanism we consider is the threat of gender-based violence. The high incidence of gender-based violence towards displaced women has been well-documented [UNHCR, 2003, Brookings Institution, 2014]. The partition of India is particularly significant in the scale and prevalence of violence against women that took place (see Bhasin and Menon [1998], Butalia [2017], Das [2006]). The salience of this risk may have induced families to marry off young women if married women were more protected by the presence of a male spouse in

et al. [2015] studies the fertility increase following an unexpected mortality shock caused by the 2004 Indonesian Tsunami). One indication that this may have occurred is that older displaced women have a higher child mortality rate than older native-born women, which is demonstrated by the fact that they have a higher number of total births but no more surviving children. If these higher child mortality rates were driven by partition, older displaced women may have had additional births to replace deceased children. This would bias the difference in differences estimates for fertility downwards, particularly for the outcome of the number of children born.

the household.

A second mechanism is the possibility of a negative wealth and income shock. Displacement is associated with the loss of physical property and assets. Additionally, displaced individuals face short and long-term income loss as they leave their preexisting sources of income and are forced to relocate to places where there might be a lack of job opportunities (e.g. refugee camps) or where their skill sets may not be well-suited (e.g. rural to urban migration). Under these circumstances, displaced parents might consider marrying their young daughters in order to alleviate financial constraints.[23]

Third, displacement events may generate changes in the composition of the marriage market that are consistent with our findings. Previous work [Bharadwaj et al., 2008] shows that the displaced population in Pakistan had higher male ratios than the non-displaced population. Consistent with our findings, the excess supply of men might have differentially increased the demand for women among the displaced group and lead to earlier marriage of displaced women. This is also corroborated by the fact that displaced women tended to marry displaced men (see Table 4).

To conclude, while all three mechanisms are consistent with our findings, we are unable to provide direct evidence on any particular mechanism. Consideration of these mechanisms, however, is important as they might imply different welfare consequences for women. The first two mechanisms point to actions that one might undertake only under extreme circumstances of violence and financial distress and it would appear to be fairly obvious that these have negative consequences for women. Yet, it is possible that women are *better off* by having married young at a time of crisis; however, by saying this we are clearly not implying that women might be better off in some overall sense, but rather that *conditional* on displacement, early marriage might have a protective aspect. The third mechanism, where marriage market changes might occur, is

[23]Corno et al. [2020] demonstrate that the sign of the response of marriage timing to negative income shocks depends on the direction of marriage payments. Traditionally, Muslim families practice bride price. In the context of Punjab, the practice of dowry is often also adopted in addition to bride price. It is unclear how much these practices were adhered to during the period of the Partition. If net payments were negative, our findings are consistent with theirs.

different as it might suggest the possibility of overall positive effects for women. If displacement expands the choice set that women have on the marriage market, or increases the demand for women, it might lead to better assortative matching or prices (through mechanisms such as a bride price or dowry) that women can command. Our results on observable characteristics of the men women married during this time does not suggest that women married "better" men; however, we wanted to note this as a possibility and an important consideration for the interpretation of these results.

While the current literature suggests that early marriage has detrimental consequences for women through lower human capital investments and higher fertility (which our results are consistent with), the question of what early marriage during a time of displacement means for women's welfare is a crucial one which, unfortunately, we are not able to answer fully due to data limitations. This is an important question for future research in this area.

Table 2.1. Summary statistics (by country)

(a) Sample composition

Country	Percent of women				Percent of men			
	Displaced	Recently displaced	Marriage migrant	Non-migrant	Displaced	Recently displaced	Marriage migrant	Non-migrant
Armenia	1.6	0		69.2	1.4	0		80.5
Cambodia	0.6	0.1	3.4	70.5	0.6	0.1	6.8	68.9
Colombia		0.8		76.1		0.9		76.8
India	0.4	0.1	31.3	60.5	0.5	0.1	0.6	88.1
Iraq	0.5	0.1	3.1	87.5	0.6	0.1	0.2	88.1
Kyrgyz Republic	0.2	0.1		68.7	0.2	0.1		76.2
Nepal	0.1	0	8	81.6	0.1	0	0.1	86.3
Pakistan	5.2			74.6	5.8			78.4
Pakistan Punjab	6.5			71.1	7.4			77.4

(b) Marriage statistics for women

Country	Average marriage age		Percentage married									
	Displaced	Non-migrant	Displaced	Non-migrant	Displaced	Non-migrant	Displaced	Non-migrant	Displaced	Non-migrant	Displaced	Non-migrant
			10 to 17		18 to 20		21 to 25		26 to 29		30 to 40	
Armenia			0.0	0.5	50.0	12.7	83.3	41.0	81.2	66.0	79.3	80.4
Cambodia	20.76	20.69	3.7	1.2	43.8	23.5	80.2	59.0	88.0	80.4	93.5	90.0
Colombia			5.4	3.6	33.8	24.9	56.6	44.4	72.4	62.5	81.0	76.5
India			20.3	5.2	81.8	39.6	91.8	71.8	98.3	90.5	99.6	97.0
Iraq	20.65	20.61	3.3	3.7	50.0	27.8	63.2	52.7	77.5	71.8	93.0	85.9
Kyrgyz Republic			0.0	0.6	44.4	22.8	61.5	61.5	90.0	84.5	90.6	94.1
Nepal	17.45	17.28	8.6	4.6	28.6	46.7	84.4	79.2	95.0	93.4	95.2	97.5
Pakistan	18.17	18.35										
Pakistan Punjab	18.42	18.63										

Note: This table reports summary statistics by country, as well as for the state of Punjab in Pakistan. We pool data for countries where we have multiple datasets. Panel (a) reports the proportion of the population classified as displaced, recently displaced (displaced in last 5 years), marriage migrants, and non-migrants by gender for each country. Panel (b) reports for average marriage ages and marriage rates by group for country. In the columns that report average marriage ages, the displaced category is comprised of women who have last migrated due to conflict or natural disaster. In the columns that report marriage rates, the displaced category is comprised of recently displaced women. Blank boxes indicate that the relevant data was not available for that country.

Table 2.2. Pakistan evidence: Age at the time of partition and marriage outcomes

(a) Young is 13 to 17

	Ever married	Child marriage	Marriage age	Married during partition
	(1)	(2)	(3)	(4)
Young X Displaced	.01***	.038**	-.28**	.053***
	(.002)	(.016)	(.126)	(.009)
Young	-.013***	.06***	-.63***	.23***
	(.001)	(.008)	(.066)	(.004)
Displaced	.00024	.019	-.25**	-.012***
	(.001)	(.013)	(.112)	(.003)
Comparison mean	.997	.308	19.3	.0169
Adjusted R^2	.00	.01	.01	.08
N	27406	26654	26654	26654

(b) Young is 1 to 5

	Ever married	Child marriage	Marriage age	Married during partition
	(1)	(2)	(3)	(4)
Young X Displaced	-.014***	-.0043	.14	.012***
	(.004)	(.016)	(.124)	(.003)
Young	-.042***	.13***	-1.2***	-.017***
	(.002)	(.008)	(.063)	(.002)
Displaced	.00024	.019	-.25**	-.012***
	(.001)	(.013)	(.112)	(.003)
Comparison mean	.997	.308	19.3	.0169
Adjusted R^2	.01	.01	.02	.01
N	35209	32967	32967	32967

Note: Each column represents a single regression estimated from Equation 2.2. The analysis sample is restricted to displaced (treatment group) and native-born (comparison group) females living in Punjab at the time of enumeration that were ages (a) 13 to 17 or 30 to 32 at the time of partition or (b) 1 to 5 or 30 to 32 at the time of partition. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 2.3. Pakistan evidence: Age at the time of partition and other own outcomes

(a) Young is 13 to 17

	Literate	Completed primary school before partition	Continued education after partition	Children born	Children surviving
	(1)	(2)	(3)	(4)	(5)
Young X Displaced	.013**	.011**	-.0027**	.43***	.55***
	(.005)	(.004)	(.001)	(.101)	(.078)
Young	.027***	.015***	.0093***	-.42***	-.22***
	(.003)	(.002)	(.001)	(.048)	(.038)
Displaced	.00026	-.0041	-3.6e-15	.34***	-.094
	(.004)	(.003)	(.)	(.087)	(.067)
Comparison mean	.0284	.0205	0	5.64	4.77
Adjusted R^2	.00	.00	.00	.01	.01
N	27406	26881	26953	27109	25657

(b) Young is 1 to 5

	Literate	Completed primary school	Children born	Children surviving
	(1)	(2)	(3)	(4)
Young X Displaced	.035***	.031***	-.052	.38***
	(.006)	(.005)	(.095)	(.074)
Young	.064***	.041***	-2.4***	-1.8***
	(.003)	(.002)	(.044)	(.035)
Displaced	.00026	-.0041	.34***	-.094
	(.004)	(.003)	(.087)	(.067)
Comparison mean	.0284	.0205	5.64	4.77
Adjusted R^2	.01	.01	.17	.15
N	35209	34110	33769	30678

Note: Each column represents a single regression estimated from Equation 2.2. The analysis sample is restricted to displaced (treatment group) and native-born (comparison group) females living in Punjab at the time of enumeration that were ages (a) 13 to 17 or 30 to 32 at the time of partition or (b) 1 to 5 or 30 to 32 at the time of partition. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 2.4. Pakistan evidence: Age at the time of partition and spousal characteristics

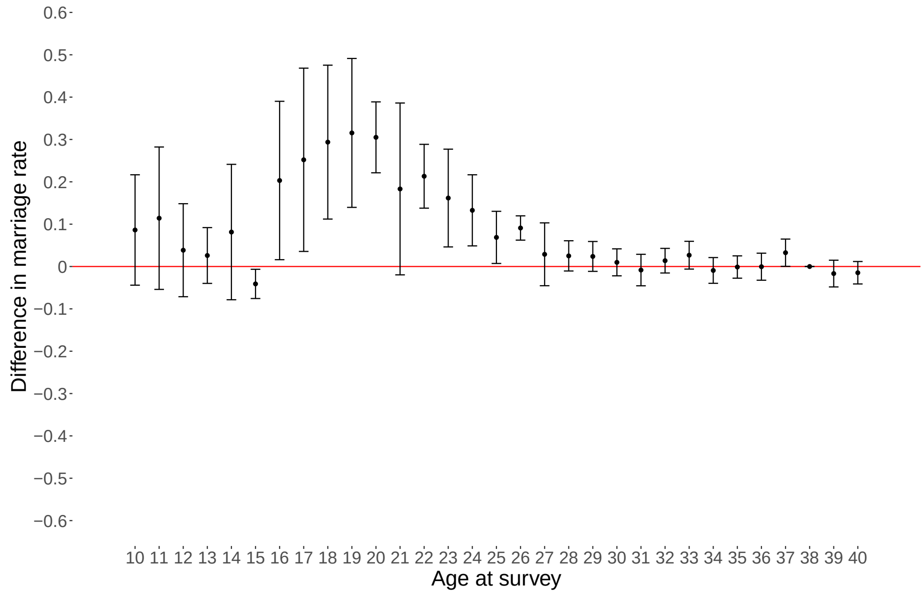
(a) Young is 13 to 17

	Age difference	Same nativity	Literate	Completed primary school	Continued education after partition
	(1)	(2)	(3)	(4)	(5)
Young X Displaced	-.13	-.0032	.0026	.0037	.00085
	(.215)	(.006)	(.014)	(.012)	(.002)
Young	.51***	-.011***	.086***	.064***	.013***
	(.107)	(.002)	(.007)	(.006)	(.001)
Displaced	.024	-.01**	.019*	.012	-.0012**
	(.173)	(.005)	(.011)	(.009)	(.001)
Comparison	6.00	.99	.15	.10	.00
Adjusted R^2	.0011	.0021	.0086	.0065	.0032
N	21807	21804	21807	20429	20950

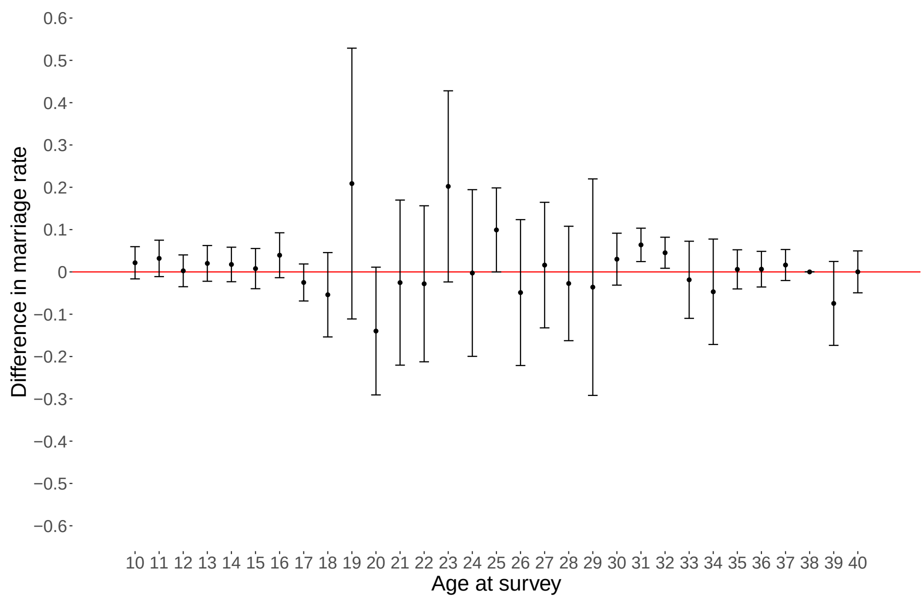
(b) Young is 1 to 5

	Age difference	Same nativity	Literate	Completed primary school	Continued education after partition
	(1)	(2)	(3)	(4)	(5)
Young X Displaced	-.58***	.011*	.054***	.057***	.065***
	(.213)	(.006)	(.015)	(.013)	(.008)
Young	-.094	-.048***	.17***	.14***	.17***
	(.101)	(.002)	(.007)	(.006)	(.003)
Displaced	.024	-.01**	.019*	.012	-.0012**
	(.173)	(.005)	(.011)	(.009)	(.001)
Comparison	6.00	.99	.15	.10	.00
Adjusted R^2	.00083	.007	.027	.024	.051
N	27390	27387	27390	25226	24930

Note: Each column represents a single regression estimated from Equation 2.2. The analysis sample is restricted to displaced (treatment group) and native-born (comparison group) females living in Punjab at the time of enumeration that were ages (a) 13 to 17 or 30 to 32 at the time of partition or (b) 1 to 5 or 30 to 32 at the time of partition. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$



(a) Women



(b) Men

Figure 2.1. Cross-country evidence: Effect of displacement on marriage rates
Note: This figure reports differences in age-specific marriage rates between recently displaced and non-migrant individuals of a given gender as indicated in each panel. Each dot represents a coefficient for δ_a estimated from Equation 2.1 with Country X Year fixed effects. The dependent variable is whether the respondent has ever been married. The error bars report a 95% for each coefficient estimated using heteroskedasticity-robust standard errors. Individual observations are weighted by inverse sampling probability.

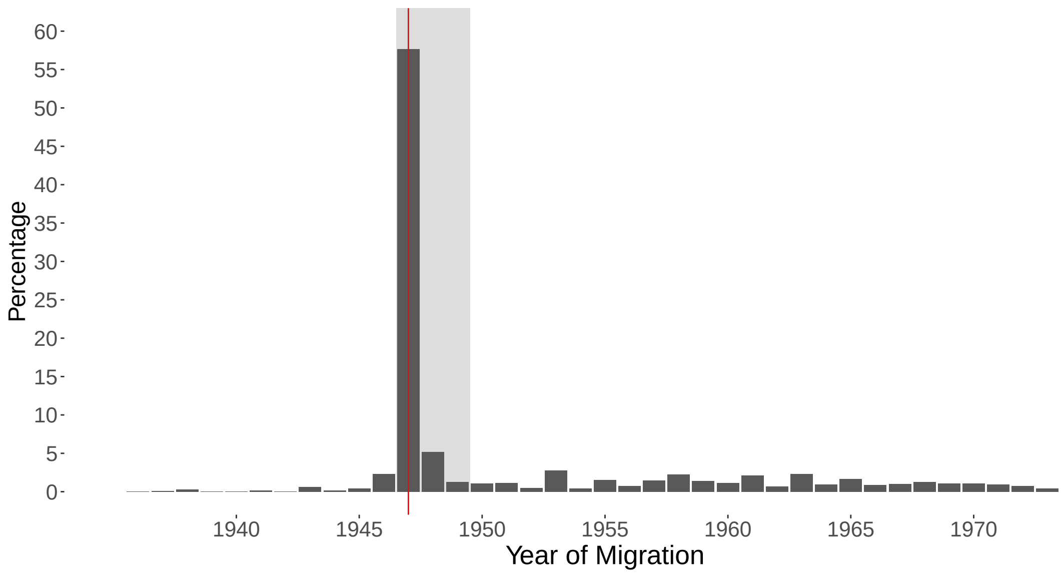
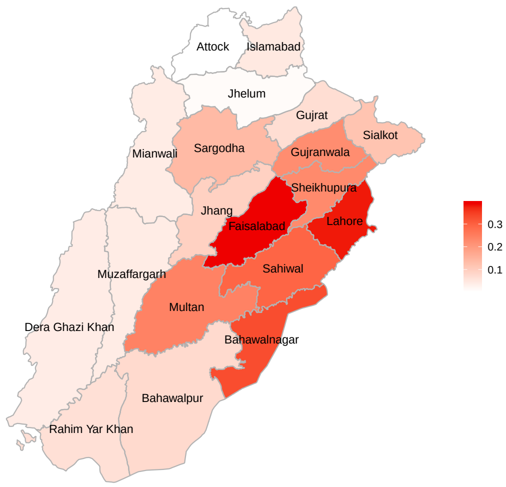
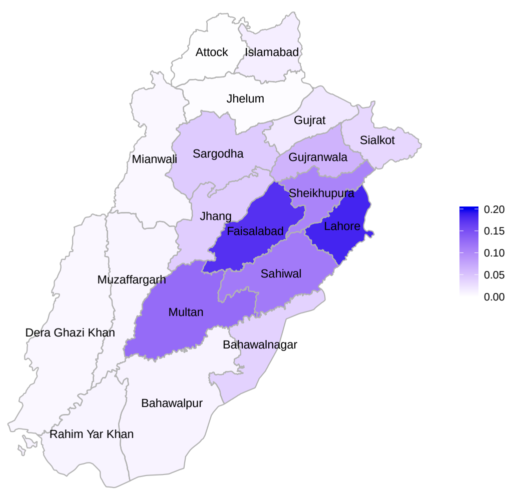


Figure 2.2. Pakistan evidence: Migration timing of Indian-born females

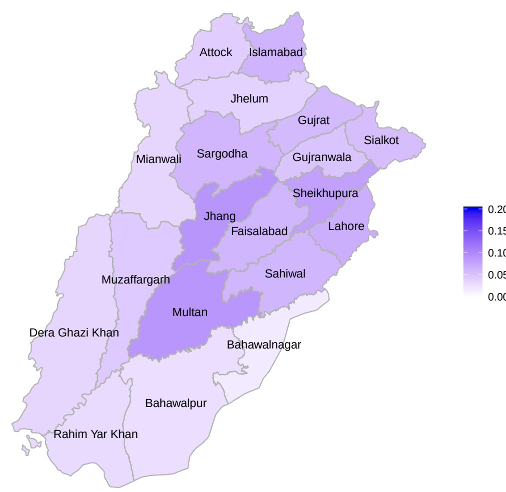
Note: This figure reports empirical densities of the year of last migration for Indian-born females by age group. The year of migration is calculated as 1973 minus the years since last migration which is reported in the data. The analysis sample is all Indian-born women alive at the time of partition who are living in Punjab at the time of the survey. The data are from the 1973 Pakistan HED survey.



(a) Proportion of women who were displaced within a given district



(b) Proportion of all displaced women in Punjab living in a given district



(c) Proportion of all native-born women in Punjab living in a given district

Figure 2.3. Pakistan evidence: Spatial distribution of Punjabi female population
Note: This figure maps the spatial distribution of native-born and displaced women living in the state of Punjab at the time of enumeration. All weighted proportions are calculated using inverse sampling probabilities. Panel (a) reports the weighted proportion of women within each district who were displaced by partition. Panel (b) reports the weighted proportion of displaced women in Punjab who were living in each district. Panel (c) reports the weighted proportion of native-born women in Punjab who were living in each district.

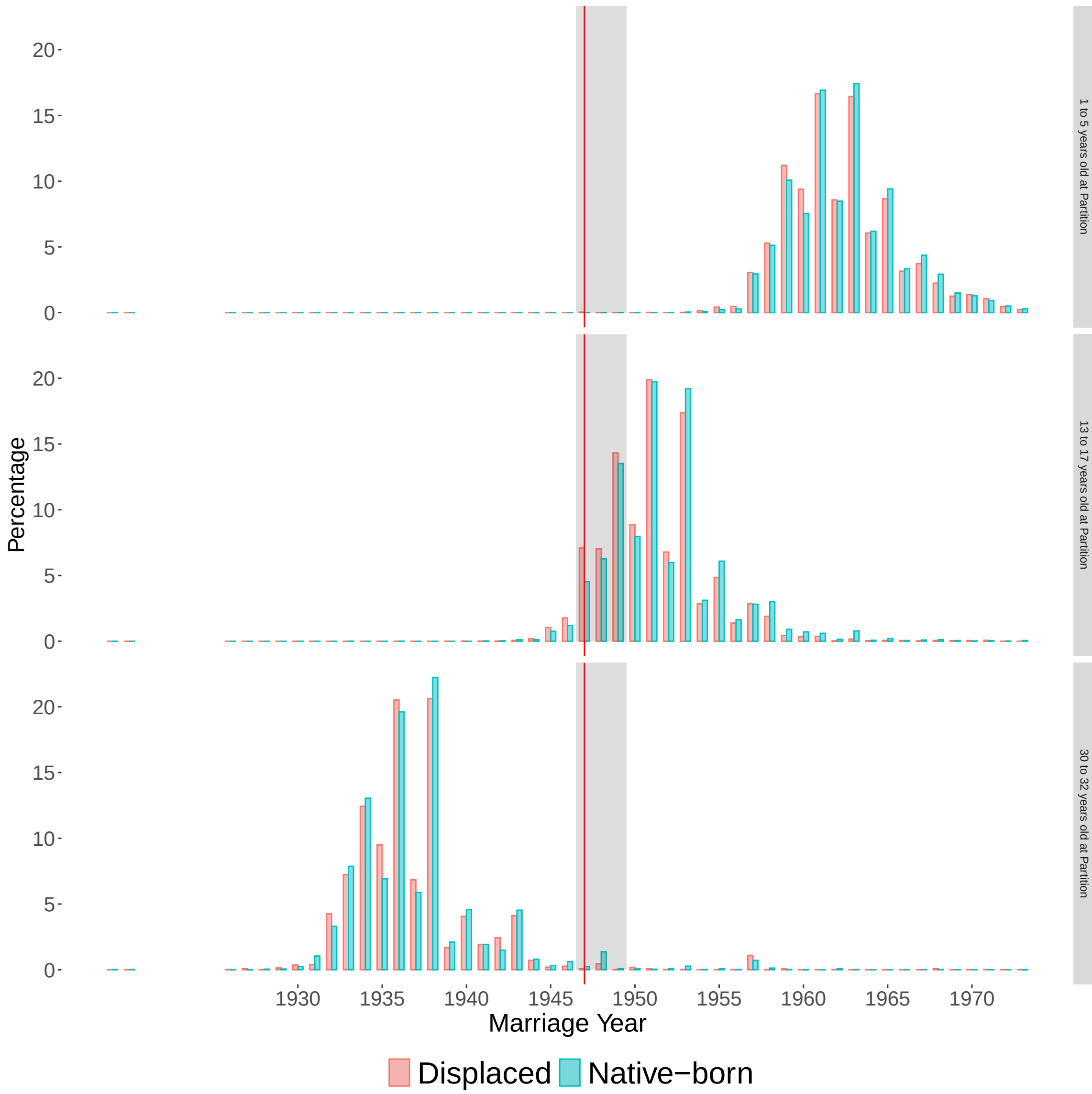


Figure 2.4. Pakistan evidence: Marriage timing by age at partition
Note: This figure plots empirical densities of the year of first marriage for native-born (born in Pakistan) and displaced (Indian-born females who migrated 24 to 26 years ago) women by age group.

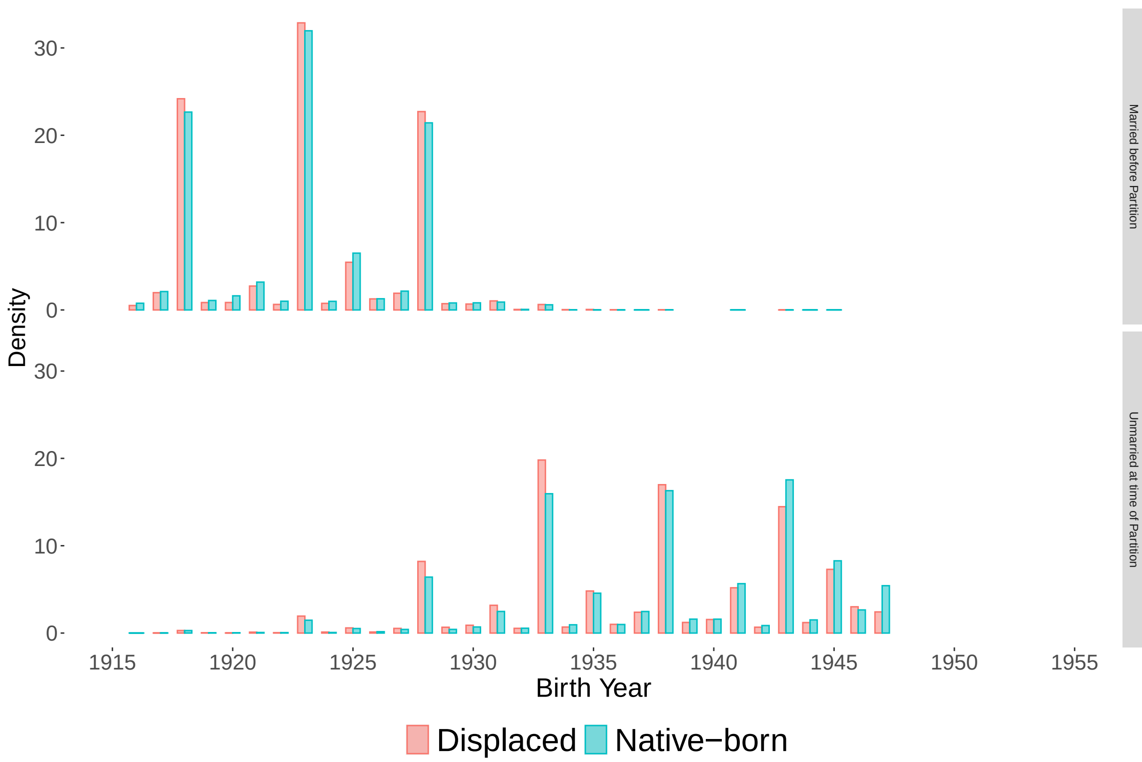
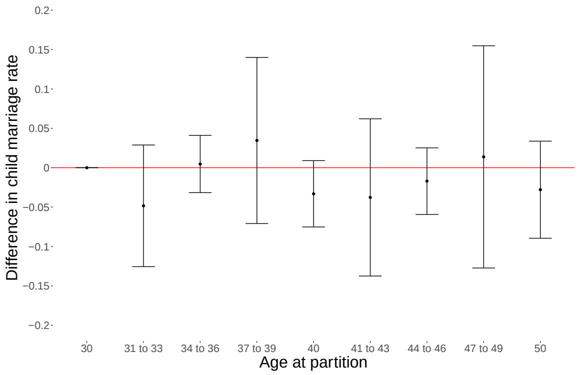
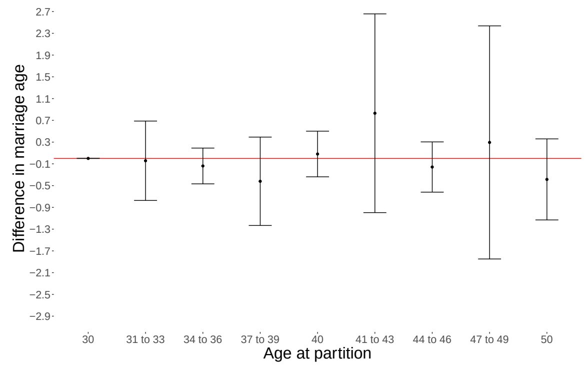


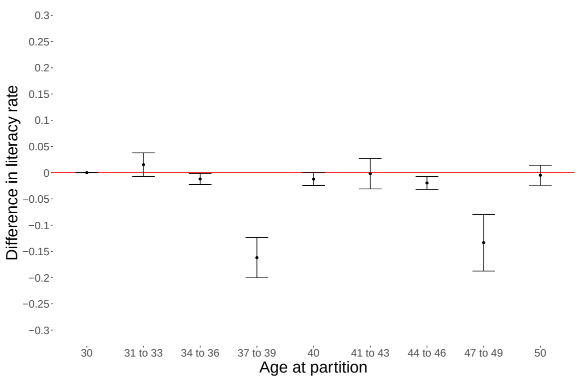
Figure 2.5. Pakistan evidence: Mortality risk at partition by marriage status
Note: This figure plots the empirical densities of the age distribution by marriage status at time of partition for displaced and native-born (control group) women.



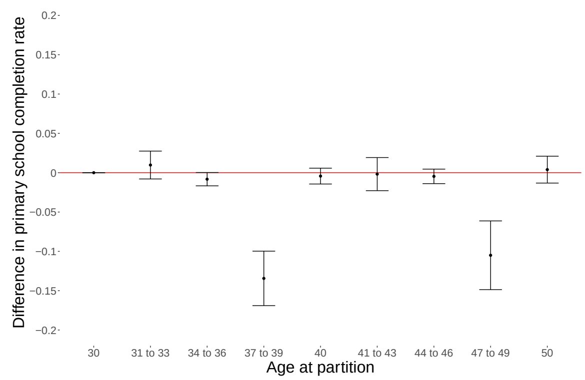
(a) Child marriage rate



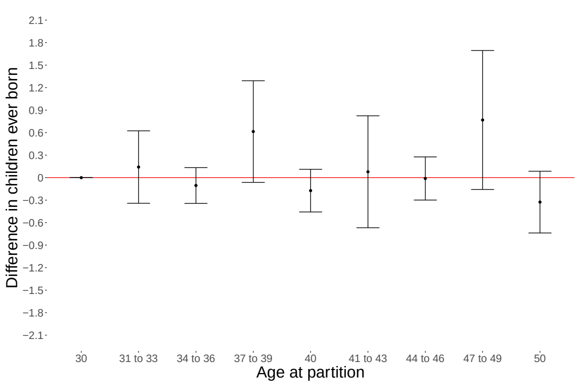
(b) Marriage age



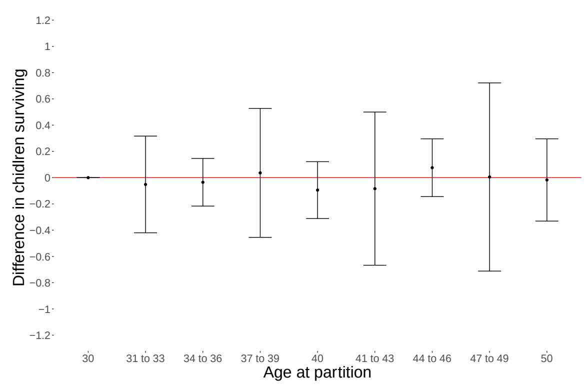
(c) Literacy



(d) Completed primary school



(e) Children ever born



(f) Children surviving

Figure 2.6. Pakistan evidence: Parallel trends

Note: This figure plots parallel trends tests for the main outcomes reported in the paper. Each dot represents the coefficient δ_a for the indicated age group a estimated from Equation 2.3. The error bars report a 95% for each coefficient estimated using heteroskedasticity-robust standard errors. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Chapter 2, in full, is a reprint of the material as it appears in The Journal of Development Economics. Lu, F., Siddiqui, S., & Bharadwaj, P. (2021). Marriage outcomes of displaced women. Journal of Development Economics, 152, 102684. The dissertation author was a primary investigator and co-author of this paper.

Chapter 3

Intergenerational Persistence in Intimate Partner Violence

3.1 Introduction

Intimate partner violence is a global human rights and public health issue. It is estimated that over one-third of women worldwide have experienced IPV [García-Moreno et al., 2013]. Intimate partner violence has been shown to negatively impact the physical and mental health of women as well as their children [Aizer, 2011, Currie et al., 2022]. In this study, we empirically explore whether and why intimate partner violence is intergenerationally persistent. We are motivated to study this question for two primary reasons. First, from a welfare perspective, we care if behavior that negatively impacts women is unequally distributed across certain families and how strong this inequality is. Second, if IPV is intergenerationally persistent, understanding why the behavior survives can help inform policies to mitigate IPV.

We estimate mother-daughter associations in intimate partner violence victimization using representative survey data from 16 countries in Sub-Saharan Africa, a region of the world that has particularly high IPV prevalence. We find women whose fathers ever physically abused their mothers are 26 percentage points (1.9x) more likely to have experienced IPV with their current cohabitating partners. This association is not explained by selection into or out of partnership.

To explain why IPV might be intergenerationally persistent, we focus on the role of cultural transmission mechanisms, which have been demonstrated to play an important role in the perpetuation of other forms of gender inequality. In particular, we focus on the formation of attitudes towards IPV and assortative matching in marriage markets. We explore whether variation in these mechanisms is across or within ethnicity and place, which are primary categories that are used to taxonomize cultural groups.

Women's attitudes towards IPV may impact their propensity to stay with abusive partners. We find that women from families with IPV are more likely to report that intimate partner violence is acceptable. Women with parent IPV have attitudes that are 0.12 σ more positive towards IPV. This difference in attitudes is consistent with family-level socialization as it does not vary across ethnicity or place. We provide evidence that this difference in attitudes is not driven by ex-post rationalization due to exposure to more violence with current partners. Assortative matching can exacerbate long-run inequality [Fernández and Rogerson, 2001, Pollak, 2004]. Overall, we find that there is assortative matching on attitudes between women and men in the partnered sample. This association is partially driven by variation across ethnicity and place. However, we find limited empirical evidence that partners of women with parent IPV have more positive attitudes towards IPV. The magnitude of this association is smaller than the product of the association between parent IPV and daughter attitudes and the association between daughter attitudes and son-in-law attitudes. This indicates that assortative matching does not play a large role in magnifying persistence.

The paper proceeds as follows. Section 3.2 describes the data and empirical methods. Section 3.3 reports and discusses our main persistence estimates. Section 3.4 discusses the role of place, ethnicity, and socioeconomic status in explaining variation in persistence. Section 3.5 presents evidence of cultural transmission mechanisms. Section 3.6 provides discussion of next steps for the paper.

3.2 Data and methods

3.2.1 Sample

Our primary data source is the Demographic and Health Surveys (DHS), which are cross-sectional representative surveys of women and men aged 15-49. We use all surveys that contain a module on domestic violence with our main variables of interest (see below). We then merge the DHS data with the Murdock Ethnographic Atlas, an ethnicity-level database containing pre-industrial cultural characteristics, using the methodology from Michalopoulos et al. [2019]. We are able to match 85% of observations from the DHS with the EA. Our main analysis sample contains 122,493 partnered women from 28 surveys in 16 countries in Sub-Saharan Africa. For a subset of 43,807 women, we are able to merge women with their partners from the men's survey. We are not able to merge all women with partners as some male partners are not present at the time of the survey, some households are not sampled for the men's survey, and only men aged 15-49 are surveyed. Table C.1 reports the list of surveys in the main and matched partner samples. The merge with the Ethnographic Atlas serves two purposes. First, the unique ethnicity identification allows us to account for the role of ethnicity separately from geography as we are able to observe the same ethnic group across different surveys. Second, it provides rich information on cultural practices that we can use to study the long-run determinants of IPV persistence.

3.2.2 Measurement

The DHS domestic violence module in the women's survey contains key information for measuring the persistence of intimate partner violence across generations. We refer to the respondent as the daughter in our context. To measure the occurrence of IPV between the respondent's parents, we use the single question of "whether the respondent's father ever beat her mother". When defining women as married, we include both formal marriages and informal unions that involve cohabitation with a partner. To measure whether a daughter has ever experienced IPV with her current husband (the "son-in-law"), we generate a single indicator variable based on a series of questions about whether the respondent has ever experienced physical or sexual violence with her current partner. Finally, both the women's and men's DHS surveys contain a set of questions on the respondent's attitudes towards domestic violence: "whether wife beating is justified" in five hypothetical situations. We combine 5 questions into a single index by calculating the first principal component across the pooled sample (standardized to mean 0 and s.d. 1).

3.2.3 Methods

In our main specification, we estimate the following linear regression for a woman i :

$$Y_i = \beta ParentIPV_i + \varepsilon_i \tag{3.1}$$

where Y_i is the outcome and $ParentIPV_i$ is an indicator variable of whether the woman reported the occurrence of IPV between her parents. β is main estimate of interest. We cluster standard errors at the DHS primary sampling unit level which are small geographic units (villages or city blocks) sampled at the largest subnational region $\times$ rural level. To explore potential mechanisms, we conduct an accounting exercise by adding covariates to the main regression specification and documenting the impact on the magnitude of the main persistence estimate β.

To investigate the role of cultural transmission and assortative matching, we correlate the attitudes on the acceptability of IPV between partnered men and women and with parent IPV. We then add increasingly granular fixed effects to control for place and ethnicity to examine whether variation in transmission is within or across cultural groups.

3.3 Pooled estimates

We report pooled estimates of intergenerational persistence in intimate partner violence in Table 3.1. The first column reports the association between a woman's reports of witnessing physical violence between her parents and whether she has ever experienced intimate partner violence with her current cohabitating partner. We find that women who report parent IPV are 25.8 percentage points (1.89x) more likely to report ever experiencing IPV with their current partner. This association broadly holds across each type of intimate partner violence: physical, sexual and emotional. Table C.2 reports the pooled estimates table for matched partner sample which are quantitatively similar.

3.3.1 Role of selection into and out of partnerships

Our persistence estimates are estimated for women who are currently partnered. One concern for interpreting these estimates is that partnership rates could vary systematically between women with and without exposure to parent intimate partner violence. In Table 3.2 we examine the role of selection into and out of partnership. First, we find that there is no difference in the proportion of women who have ever been in a partnership (cohabitation or marriage) by parent IPV (Column 1). However, women with parent IPV are slightly less likely to be currently partnered (Column 2), which is the selection criteria for our analysis sample. This is driven primarily by higher rates of divorce and separation, but higher rates of widowship also play a role (Columns 5 and 6).

We then conduct a Lee Bounds exercise to determine whether selection out of partnership quantitatively impacts our persistence estimates. Table C.3 reports these Lee Bounds for Table C.2. We estimate very tight bounds for our persistence estimates, indicating that they are not driven by this selection.

3.4 Account for persistence

We examine the role of place and ethnicity in explaining intergenerational persistence in IPV could be explained by persistent spatial inequality and cultural factors. To examine the role of place, we cumulatively increase the granularity of the controls from the baseline specification to include survey (which is almost equivalent to country) and local region (the largest subnational region interacted with rural) fixed effects. Place-based effects at the regional level explain 33% of variation persistence. Next, we control for ethnicity. Ethnicity fixed effects alone account for 32 % of persistence variation. However, ethnic groups are highly spatially clustered, so we examine the combined role of place and ethnicity and find that the interaction of ethnicity and local region together explain around 40% of persistence variation.

A majority of the intergenerational persistence in IPV is driven by within locality and within ethnicity variation. One potential explanation is that socioeconomic status is an important determinant for the incidence of intimate partner violence and is also intergenerationally persistent. To explore this, we control for measures of socioeconomic status including the respondent's education, her husband's education and their current household wealth. We find that this accounts for almost none of the variation in persistence. This allows us to further control for the male partner's ethnicity and we find that it doesn't play a quantitatively significant role on top of the female respondent's ethnicity.

3.5 Cultural transmission

We examine the potential role of cultural transmission mechanisms in the intergenerational persistence of intimate partner violence. First, we study the formation of attitudes towards IPV. Variation in attitudes can be social group based due to ethnicity, place based due to social norms or peer effects or, family based due to the parent socialization of exposed daughters. Second, we study the role of assortative matching on in the marriage market. Variation in assortative matching can be due to traditional marriage practices, place-based due to local marriage market conditions, or family-based due to preferences for marrying people "similar".

We use data from the matched partner sample, which contains a data for partnered women and men on the acceptability of intimate partner violence. The module contains 5 question about whether "wife beating" is justified in a given hypothetical situation. We use principal component analysis to create an index using the first principal component and standardize the index to have mean 0 and variance 1. To study the formation of attitudes, we correlate daughter attitudes with parent IPV. To study assortative matching, we correlate son-in-law attitudes with daughter attitudes and parent IPV.

Table 3.3 reports the relationship between daughter attitudes and parent IPV. Exposure to parent IPV is associated with a .107 σ increase in the daughter's reported acceptability of intimate partner violence (Column 7). These results are quantitatively similar to those estimated on the whole sample (Column 1). We find that this positive association in attitudes is even stronger for unpartnered women (Column 4), which provides evidence against ex-post rationalization. Additionally, we find that ethnicity and place do not explain the relationship between parent IPV and daughter attitudes. This is consistent with direct exposure to parent IPV playing a role in the socialization of daughters.

Table 3.4a reports the relationship between son-in-law attitudes with daughter attitudes and parent attitudes. Son-in-law attitudes have a strong positive association with daughter attitudes. A one standard deviation increase in daughter attitudes is associated with a 0.156 σ increase in son-in-law attitudes. A large portion of this association is explained by ethnicity and place-level variation. However, there is not a strong relationship between son-in-law attitudes and the daughter's exposure to intimate violence between her parents. Within ethnicity, there is a weak positive association that is partially explained by place. The magnitude of this association is smaller than the product of the association between parent IPV and daughter attitudes and the association between daughter attitudes and son-in-law attitudes. This indicates that assortative matching does not play a large role in magnifying persistence.

3.6 Discussion

In the next steps of the project, we plan to explore the role of cultural practices and norms. First, we will examine heterogeneity by marriage practices, gender norms, and family structure using ethnicity-level variation from the Ethnographic Atlas. Second, using census data and ethnic homelands data from the Murdock Map, we will exploit geographic variation in whether the respondent's ethnic group is a local minority group, as this will determine the size of the co-ethnic marriage market and incentives for homogamy [Bisin and Verdier, 2000]. Finally, we plan to quantify the potential role of respondent reporting error and bias in explaining our findings using Monte Carlo simulations.

Table 3.1. Pooled estimates

	X violence with current partner?			
	(1) Any	(2) Physical	(3) Sexual	(4) Emotional
Parent IPV	.258***	.232***	.109***	.167***
	[.005]	[.0048]	[.0036]	[.0048]
Comparison Mean	.29	.18	.07	.20
Observations	122493	122493	122493	122493
R^2	.05	.05	.02	.03

Note: Each column represents a single regression estimated from Equation 3.1. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 3.2. Selection into and out of partnership

	X partnered?				Unpartnered due to X?	
	(1) Ever	(2) Current	(3) # times	(4) Age first	(5) Divorce/separation	(6) Widowed
Parent IPV	.002	-.029***	.031***	.206***	.023***	.008***
	[.004]	[.0044]	[.0034]	[.042]	[.0021]	[.0014]
Comparison Mean	.82	.75	1.14	18.05	.05	.03
Observations	165619	165619	135683	136103	165619	165619
R^2	.00	.00	.00	.00	.00	.00

Note: Each column represents a single regression estimated from Equation 3.1. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 3.3. Daughter's attitudes towards IPV

	Whole sample			Unpartnered sample			Matched partner sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Parent IPV	.112***	.127***	.127***	.266***	.231***	.198***	.107***	.149***	.123***	.123***
	[.012]	[.0097]	[.0097]	[.02]	[.019]	[.018]	[.018]	[.016]	[.015]	[.015]
Observations	119531	119531	119531	26626	26626	26626	43807	43807	43807	43807
R^2	.00	.25	.25	.02	.13	.25	.00	.16	.26	.26
Daughter ethnicity FEs	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Daughter ethnicity x Region FEs	No	No	Yes	No	No	Yes	No	No	Yes	Yes
Son-in-law ethnicity FEs	No	No	No	No	No	No	No	No	No	Yes

Table 3.4. Son-in-law attitudes towards IPV

(a) Relationship with daughter attitudes

	Std. index of son-in-law's acceptance of IPV			
	(1)	(2)	(3)	(4)
Daughter's acceptability	.156***	.093***	.061***	.062***
	[.0078]	[.0076]	[.0077]	[.0077]
Observations	43807	43807	43807	43807
R^2	.03	.09	.16	.16
Daughter ethnicity FEs	No	Yes	Yes	Yes
Daughter ethnicity x Region FEs	No	No	Yes	Yes
Son-in-law ethnicity FEs	No	No	No	Yes

(b) Relationship with parent IPV

	Std. index of son-in-law's acceptance of IPV			
	(1)	(2)	(3)	(4)
Daughter's parent IPV	.002	.036*	.017	.019
	[.016]	[.015]	[.014]	[.014]
Observations	43807	43807	43807	43807
R^2	.00	.08	.16	.16
Daughter ethnicity FEs	No	Yes	Yes	Yes
Daughter ethnicity x Region FEs	No	No	Yes	Yes
Son-in-law ethnicity FEs	No	No	No	Yes

Note: Each column represents a single regression estimated from Equation 3.1. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Bibliography

Richard Gyan Aboagye, Abdul-Aziz Seidu, Bernard Yeboah-Asiamah Asare, Prince Peprah, Isaac Yeboah Addo, and Bright Opoku Ahinkorah. Exposure to interparental violence and justification of intimate partner violence among women in sexual unions in sub-saharan africa. *Archives of Public Health*, 79(1):1–11, 2021.

Anna Aizer. Poverty, violence, and health the impact of domestic violence during pregnancy on newborn health. *Journal of Human resources*, 46(3):518–538, 2011.

Diego Alburez-Gutierrez, Martin Kolk, and Emilio Zagheni. Women's experience of child death over the life course: A global demographic perspective. *Demography*, 2021.

Alberto Alesina, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou. Intergenerational mobility in africa. *Econometrica*, 89(1):1–35, 2021.

Sam Asher, Paul Novosad, and Charlie Rafkin. Intergenerational mobility in india: New methods and estimates across time, space, and communities, 2020.

Anthony Atkinson and Francois Bourguignon. The comparison of multi-dimensioned distributions of economic status. *The Review of Economic Studies*, 49(2):183–201, 1982.

Mirza Balaj, Hunter Wade York, Kam Sripada, Elodie Besnier, Hanne Dahl Vonen, Aleksandr Aravkin, Joseph Friedman, Max Griswold, Magnus Rom Jensen, and Talal Mohammad. Parental education and inequalities in child mortality: A global systematic review and meta-analysis. *Lancet*, 398(10300):608–620, 2021.

Jean-Marie Baland, Guilhem Cassan, and Benoit Decerf. 'too young to die': Deprivation measures combining poverty and premature mortality. *American Economic Journal: Applied Economics*, 13(4):226–57, October 2021.

Abhijit Banerjee and Andrew Newman. Occupational choice and the process of development. *Journal of Political Economy*, 101(2):274–298, 1993.

Gary S Becker, Tomas J Philipson, and Rodrigo R Soares. The quantity and quality of life and the evolution of world inequality. *American Economic Review*, 95(1):277–291, 2005.

Sascha O Becker and Andreas Ferrara. Consequences of forced migration: A survey of recent findings. *Labour Economics*, 59:1–16, 2019.

Sascha O Becker, Irena Grosfeld, Pauline Grosjean, Nico Voigtlander, and Ekaterina Zhuravskaya. Forced migration and human capital: Evidence from post-wwii population transfers. *American Economic Review*, 110(5):1430–63, 2020.

Sonia Bhalotra and Samantha Rawlings. Gradients of the intergenerational transmission of health in developing countries. *Review of Economics and Statistics*, 95(2):660–672, 2013.

Prashant Bharadwaj and James Fenske. Partition, migration, and jute cultivation in india. *Journal of Development Studies*, 48(8):1084–1107, 2012.

Prashant Bharadwaj and Rinchan Ali Mirza. Displacement and development: Long term impacts of population transfer in india. *Explorations in Economic History*, 73, 2019.

Prashant Bharadwaj, Asim Khwaja, and Atif Mian. The big march: migratory flows after the partition of india. *Economic and Political Weekly*, pages 39–49, 2008.

Kamla Bhasin and Ritu Menon. Borders and boundaries: women in india's partition. *New Delhi: Kali for Women*, 1998.

Manudeep Bhuller, Gordon B Dahl, Katrine V Løken, and Magne Mogstad. Domestic violence and the mental health and well-being of victims and their children. Technical report, National Bureau of Economic Research, 2022.

Alberto Bisin and Thierry Verdier. "beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics*, 115(3): 955–988, 2000.

Sandra Black, Neil Duzett, Adriana Lleras-Muney, Nolan Pope, and Joseph Price. Intergenerational correlations in longevity, 2022.

Sandra E Black, Paul J Devereux, and Kjell G Salvanes. From the cradle to the labor market? the effect of birth weight on adult outcomes. *Quarterly Journal of Economics*, 122(1):409–439, 2007.

John Bound, Charles Brown, and Nancy Mathiowetz. Measurement error in survey data. In James Heckman and Edward Leamer, editors, *Handbook of Econometrics*, volume 5, pages 3705–3843. North-Holland, Amsterdam, 2001.

Megan Bradley. Refugee repatriation: justice, responsibility and redress. 2013.

Brookings Institution. Improving the protection of internally displaced women: Assessment of progress and challenges. Technical report, 2014.

Roy Burstein, Nathaniel J Henry, Michael L Collison, Laurie B Marczak, Amber Sligar, Stefanie Watson, Neal Marquez, Mahdieh Abbasalizad-Farhangi, Masoumeh Abbasi, Foad Abd-Allah, et al. Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature*, 574(7778):353–358, 2019.

Urvashi Butalia. *The other side of silence: Voices from the partition of India*. Penguin UK, 2017.

John C Caldwell. Routes to low mortality in poor countries. *Population and Development Review*, pages 171–220, 1986.

A Colin Cameron, Jonah B Gelbach, and Douglas L Miller. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427, 2008.

Ivan A Canay, Andres Santos, and Azeem M Shaikh. The wild bootstrap with a "small" number of "large" clusters. *The Review of Economics and Statistics*, 2021.

Francesco Caselli. Accounting for cross-country income differences. *Handbook of economic growth*, 1:679–741, 2005.

AV Chari, Rachel Heath, Annemie Maertens, and Freeha Fatima. The causal effect of maternal age at marriage on child wellbeing: Evidence from india. *Journal of Development Economics*, 127:42–55, 2017.

Madan G Chitkara. *Converts do not make a nation*. APH Publishing, 1998.

Bernice H Cohen. Family patterns of mortality and life span. *The Quarterly Review of Biology*, 39(2):130–181, 1964.

Carolina VN Coll, Thiago M Santos, Karen Devries, Felicia Knaul, Flavia Bustreo, Anne Gatuguta, Gbenankpon Mathias Houvessou, and Aluísio JD Barros. Identifying the women most vulnerable to intimate partner violence: A decision tree analysis from 48 low and middle-income countries. *EClinicalMedicine*, 42:101214, 2021.

Lucia Corno, Nicole Hildebrandt, and Alessandra Voena. Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica*, 88(3):879–915, 2020.

Janet Currie and Enrico Moretti. Biology as destiny? short-and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor Economics*, 25(2):231–264, 2007.

Janet Currie and Tom Vogl. Early-life health and adult circumstance in developing countries. *Annual Review of Economics*, 5(1):1–36, 2013.

Janet Currie, Michael Mueller-Smith, and Maya Rossin-Slater. Violence while in utero: The impact of assaults during pregnancy on birth outcomes. *Review of Economics and Statistics*, 104(3):525–540, 2022.

Veena Das. *Life and Words: Violence and the Descent into the Ordinary*. Univ of California Press, 2006.

Ori Davidov, David Faraggi, and Benjamin Reiser. Misclassification in logistic regression with discrete covariates. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 45(5):541–553, 2003.

Richard de Groot, Maxwell Yiryele Kuunyem, and Tia Palermo. Child marriage and associated outcomes in northern ghana: a cross-sectional study. *BMC public health*, 18(1):285, 2018.

Marcos Delprato, Kwame Akyeampong, and Máiréad Dunne. Intergenerational education effects of early marriage in sub-saharan africa. *World Development*, 91:173–192, 2017.

Nata Duvvury, Aoife Callan, Patrick Carney, and Srinivas Raghavendra. Intimate partner violence: Economic costs and implications for growth and development. 2013.

F Fasani, T Frattini, and L Minale. (the struggle for) refugee integration into the labour market: Evidence from europe. *CEPR Discussion Paper No. DP12718*, 2018.

Robert C Feenstra, Robert Inklaar, and Marcel P Timmer. The next generation of the penn world table. *American Economic Review, 105(10), 3150-3182, available for download at www.ggdc.net/pwt*, 2015.

Raquel Fernández and Richard Rogerson. Sorting and long-run inequality. *The Quarterly Journal of Economics*, 116(4):1305–1341, 2001.

Erica Field and Attila Ambrus. Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of political Economy*, 116(5):881–930, 2008.

Marc Fleurbaey. Beyond gdp: The quest for a measure of social welfare. *Journal of Economic Literature*, 47(4):1029–75, 2009.

Marc Fleurbaey and Guillaume Gaulier. International comparisons of living standards by equivalent incomes. *Scandinavian Journal of Economics*, 111(3):597–624, 2009.

Anita J Gagnon, Lisa Merry, and Cathlyn Robinson. A systematic review of refugee women's reproductive health. *Refuge: Canada's Journal on Refugees*, pages 6–17, 2002.

Oded Galor and Joseph Zeira. Income distribution and macroeconomics. *Review of Economic Studies*, 60(1):35–52, 1993.

Claudia García-Moreno, Christina Pallitto, Karen Devries, Heidi Stöckl, Charlotte Watts, and Naeema Abrahams. *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*. World Health Organization, 2013.

Nils Petter Gleditsch, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. Armed conflict 1946-2001: A new dataset. *Journal of peace research*, 39(5):615–637, 2002.

Srinivas Goli, Deepti Singh, and TV Sekher. Exploring the myth of mixed marriages in india: evidence from a nation-wide survey. *Journal of comparative family studies*, 44(2):193–206, 2013.

Timothy J Halliday, Bhashkar Mazumder, and Ashley Wong. Intergenerational mobility in self-reported health status in the us. *Journal of Public Economics*, 193:104307, 2021.

Stéphane Helleringer, Gilles Pison, Almamy M Kanté, Géraldine Duthé, and Armelle Andro. Reporting errors in siblings' survival histories and their impact on adult mortality estimates: Results from a record linkage study in senegal. *Demography*, 51(2):387–411, 2014.

Anne Maria Herskind, Matthew McGue, Niels V Holm, Thorkild IA Sörensen, Bent Harvald, and James W Vaupel. The heritability of human longevity: A population-based study of 2872 danish twin pairs born 1870–1900. *Human Genetics*, 97(3):319–323, 1996.

Ivan Iachine, Axel Skytthe, James W Vaupel, Matt McGue, Markku Koskenvuo, Jaakko Kaprio, Nancy L Pedersen, Kaare Christensen, et al. Genetic influence on human lifespan and longevity. *Human Genetics*, 119(3):312–321, 2006.

ICF. Demographic and health surveys (various) [datasets]. *Funded by USAID. Rockville, Maryland: ICF [Distributor]*, 2004-2017.

iDMC. Global Disaster Displacement Risk: A baseline for future work. Technical report, 2017.

Christophe Jaffrelot. *The Pakistan paradox: Instability and resilience*. Oxford University Press, 2015.

Ayesha Jalal. *The sole spokesman: Jinnah, the Muslim League and the demand for Pakistan*, volume 31. Cambridge University Press, 1994.

Lawrence James. *Raj: the making and unmaking of British India*. Hachette UK, 2010.

Seema Jayachandran. The roots of gender inequality in developing countries. *economics*, 7(1): 63–88, 2015.

Robert Jensen and Rebecca Thornton. The consequences of early marriage in the developing world. *Oxfam Journal of Gender and Development*, 11(2):9–19, 2003.

Saumitra Jha and Steven Wilkinson. Does combat experience foster organizational skill? evidence from ethnic cleansing during the partition of south asia. *American Political Science Review*, 106(4):883–907, 2012.

Charles I Jones and Peter J Klenow. Beyond gdp? welfare across countries and time. *American Economic Review*, 106(9):2426–57, 2016.

Ines Keygnaert, Nicole Vettenburg, and Marleen Temmerman. Hidden violence is silent rape: sexual and gender-based violence in refugees, asylum seekers and undocumented migrants in belgium and the netherlands. *Culture, health & sexuality*, 14(5):505–520, 2012.

Omar Khalidi. From torrent to trickle: Indian muslim migration to pakistan, 1947—97. *Islamic studies*, 37(3):339–352, 1998.

Yasmin Khan. *The great partition: The making of India and Pakistan*. Yale University Press, 2017.

Serge-Christophe Kolm. Multidimensional egalitarianisms. *Quarterly Journal of Economics*, pages 1–13, 1977.

Alan Krueger. The rise and consequences of inequality in the united states, 2012. Speech given at Center for American Progress, Washington, DC, January 12.

Simon Kuznets. Economic growth and income inequality. *American Economic Review*, 45(1): 1–28, 1955.

Jiong Li, Dorthe Hansen Precht, Preben Bo Mortensen, and Jørn Olsen. Mortality in parents after death of a child in denmark: A nationwide follow-up study. *Lancet*, 361(9355):363–367, 2003.

Jiong Li, Thomas Munk Laursen, Dorthe Hansen Precht, Jørn Olsen, and Preben Bo Mortensen. Hospitalization for mental illness among parents after the death of a child. *New England Journal of Medicine*, 352(12):1190–1196, 2005.

Frances Lu and Tom Vogl. Code for: Intergenerational persistence in child mortality. *American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/E163801V1*, 2022.

Aprajit Mahajan. Identification and estimation of regression models with misclassification. *Econometrica*, 74(3):631–665, 2006.

Maia Maia, Nasma Berri, and Sawsan Abdulrahim. Marriage and fertility patterns among jordanians and syrian refugees in jordan. Economic Research Forum, 2018.

Momoe Makino. Marriage, dowry, and women's status in rural punjab, pakistan. *Journal of Population Economics*, 32(3):769–797, 2019.

Bruno Masquelier and Catriona Dutreuilh. Sibship sizes and family sizes in survey data used to estimate mortality. *Population*, 69(2):221–238, 2014.

Amelia Reese Masterson, Jinan Usta, Jhumka Gupta, and Adrienne S Ettinger. Assessment of reproductive health and violence against women among displaced syrians in lebanon. *BMC women's health*, 14(1):25, 2014.

Stelios Michalopoulos, Louis Putterman, and David N Weil. The influence of ancestral lifeways on individual economic outcomes in sub-saharan africa. *Journal of the European Economic Association*, 17(4):1186–1231, 2019.

Minnesota Population Center. Integrated Public Use Microdata Series, 2019.

Dilip Mookherjee and Debraj Ray. Persistent inequality. *Review of Economic Studies*, 70(2): 369–393, 2003.

Meghan D Morris, Steve T Popper, Timothy C Rodwell, Stephanie K Brodine, and Kimberly C Brouwer. Healthcare barriers of refugees post-resettlement. *Journal of community health*, 34 (6):529, 2009.

Rima Mourtada, Jennifer Schlecht, and Jocelyn DeJong. A qualitative study exploring child marriage practices among syrian conflict-affected populations in lebanon. *Conflict and health*, 11(1):27, 2017.

Ercio Muñoz. The geography of intergenerational mobility in latin america and the caribbean, 2021.

Ambar Narayan, Roy Van der Weide, Alexandru Cojocaru, Christoph Lakner, Silvia Redaelli, Daniel Gerszon Mahler, Rakesh Ramasubbaiah, and Stefan Thewissen. *Fair Progress? Economic Mobility Across Generations Around the World. Equity and Development*. World Bank, Washington, DC, 2018.

Guido Neidhöfer, Joaquín Serrano, and Leonardo Gasparini. Educational inequality and intergenerational mobility in latin america: A new natabase. *Journal of Development Economics*, 134:329–349, 2018.

Jenna Nobles, Elizabeth Frankenberg, and Duncan Thomas. The effects of mortality on fertility: population dynamics after a natural disaster. *Demography*, 52(1):15–38, 2015.

Nawal M Nour. Child marriage: a silent health and human rights issue. *Reviews in obstetrics and gynecology*, 2(1):51, 2009.

Valerie Kincade Oppenheimer. A theory of marriage timing. *American journal of sociology*, 94 (3):563–591, 1988.

Therese Pettersson, Shawn Davies, Amber Deniz, Garoun Engström, Nanar Hawach, Stina Högbladh, and Margareta Sollenberg Magnus Öberg. Organized violence 1989–2020, with a special emphasis on syria. *Journal of Peace Research*, 58(4):809–825, 2021.

Thomas Piketty. Theories of persistent inequality and intergenerational mobility. *Handbook of Income Distribution*, 1:429–476, 2000.

Robert A Pollak. An intergenerational model of domestic violence. *Journal of Population Economics*, 17(2):311–329, 2004.

Anita Raj, Marissa Salazar, Emma C Jackson, Natalie Wyss, Katherine A McClendon, Aarushi Khanna, Yemeserach Belayneh, and Lotus McDougal. Students and brides: a qualitative analysis of the relationship between girl's education and early marriage in ethiopia and india. *BMC public health*, 19(1):1–20, 2019.

Catherine Rogers, Frank Floyd, Marsha Mailick Seltzer, Jan Greenberg, and Jinkuk Hong. Long-term effects of the death of a child on parents' adjustment in midlife. *Journal of Family Psychology*, 22(2):203, 2008.

Heather Royer. Separated at girth: Us twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, 1(1):49–85, 2009.

Jennifer Schlecht, Elizabeth Rowley, and Juliet Babirye. Early relationships and marriage in conflict and post-conflict settings: vulnerability of youth in uganda. *Reproductive Health Matters*, 21(41):234–242, 2013.

Sheetal Sekhri and Sisir Debnath. Intergenerational consequences of early age marriages of girls: Effect on children's human capital. *The Journal of Development Studies*, 50(12):1670–1686, 2014.

Amartya Sen. *Development as Freedom*. Oxford University Press, Oxford, 1999.

Nasra M Shah, JT Fawcett, SE Khoo, and PC Smith. The female migrant in pakistan. *Women In The Cities Of Asia: Migration And Urban Adaptation*, Chapter 6, 1983.

Maia Sieverding, Caroline Krafft, Nasma Berri, Caitlyn Keo, Mariam Sharpless, et al. Education interrupted: Enrollment, attainment, and dropout of syrian refugees in jordan. In *Economic Research Forum Working Paper Series No*, volume 1257, 2018.

Selcuk R Sirin and Lauren Rogers-Sirin. *The educational and mental health needs of Syrian refugee children*. Migration Policy Institute Washington, DC, 2015.

Emily Smith-Greenaway and Jenny Trinitapoli. Maternal cumulative prevalence measures of child mortality show heavy burden in sub-saharan africa. *Proceedings of the National Academy of Sciences*, 117(8):4027–4033, 2020.

Emily Smith-Greenaway, Diego Alburez-Gutierrez, Jenny Trinitapoli, and Emilio Zagheni. Global burden of maternal bereavement: Indicators of the cumulative prevalence of child loss. *BMJ Global Health*, 6(4), 2021.

Abbey Steele et al. Massive civilian displacement in civil war: Assessing variation in colombia. Technical report, Households in Conflict Network, 2007.

Margaret Stroebe, Henk Schut, and Wolfgang Stroebe. Health outcomes of bereavement. *Lancet*, 370(9603):1960–1973, 2007.

Ian Talbot and Gurharpal Singh. *The partition of India*. Cambridge University Press Cambridge, 2009.

Tai Yong Tan, Gyanesh Kudaisya, et al. *The aftermath of partition in South Asia*, volume 3. Psychology Press, 2000.

Florencia Torche. Intergenerational mobility and inequality: The latin american case. *Annual Review of Sociology*, 40:619–642, 2014.

James Trussell and German Rodriguez. A note on the sisterhood estimator of maternal mortality. *Studies in Family Planning*, 21(6):344–346, 1990.

Debra Umberson, Julie Skalamera Olson, Robert Crosnoe, Hui Liu, Tetyana Pudrovska, and Rachel Donnelly. Death of family members as an overlooked source of racial disadvantage in the united states. *Proceedings of the National Academy of Sciences*, 114(5):915–920, 2017.

UNHCR. Sexual and gender-based violence against refugees, returnees and internally displaced persons-guidelines for prevention and response, 2003.

UNHCR. Global Trends: Forced Displacement in 2018. Technical report, 2019. URL https://www.unhcr.org/globaltrends2018/{#}.

United Nations. *World Population Prospects 2019*. United Nations, Department of Economic and Social Affairs, Population Division, New York, 2019.

Department of Economic United Nations and Population Division Social Affairs. World population prospects 2019: Data booklet (st/esa/ser. a/424), 2019.

Jinan Usta and Amelia Reese Masterson. Women and health in refugee settings: The case of displaced syrian women in lebanon. In *Gender-based violence*, pages 119–143. Springer, 2015.

Jinan Usta, Jo Ann M Farver, and Lama Zein. Women, war, and violence: surviving the experience. *Journal of Women's Health*, 17(5):793–804, 2008.

James W Vaupel. Inherited frailty and longevity. *Demography*, 25(2):277–287, 1988.

Atheendar S Venkataramani. The intergenerational transmission of height: Evidence from rural vietnam. *Health Economics*, 20(12):1448–1467, 2011.

Tom Vogl. Dynamics of child mortality inequality. *AEA Papers and Proceedings*, 108:348–52, 2018.

Tom S Vogl. Marriage institutions and sibling competition: evidence from south asia. *The Quarterly Journal of Economics*, 128(3):1017 – 1072, 2013.

Alexander Vu, Atif Adam, Andrea Wirtz, Kiemanh Pham, Leonard Rubenstein, Nancy Glass, Chris Beyrer, and Sonal Singh. The prevalence of sexual violence among female refugees in complex humanitarian emergencies: a systematic review and meta-analysis. *PLoS currents*, 6, 2014.

Andrea L Wirtz, Kiemanh Pham, Nancy Glass, Saskia Loochkartt, Teemar Kidane, Decssy Cuspoca, Leonard S Rubenstein, Sonal Singh, and Alexander Vu. Gender-based violence in conflict and displacement: qualitative findings from displaced women in colombia. *Conflict and health*, 8(1):10, 2014.

Alexa R Yakubovich, Heidi Stöckl, Joseph Murray, GJ Melendez-Torres, Janina I Steinert, Calla EY Glavin, and David K Humphreys. Risk and protective factors for intimate partner violence against women: Systematic review and meta-analyses of prospective–longitudinal studies. *American journal of public health*, 108(7):e1–e11, 2018.

Philip Ziegler. *Mountbatten: the official biography*. Fontana Press, 1985.

Solly Zuckerman. Earl mountbatten of burma, kg, om, 25 june 1900-27 august 1979, 1981.

Appendix A for Chapter 1

Table A.1. Demographic and Health Surveys in the Sample

Afghanistan: 2010, 2015	Lesotho: 2004, 2009, 2014
Bangladesh: 2001	Madagascar: 1992, 1997, 2004, 2009
Benin: 1996, 2008	Malawi: 1992, 2000, 2004, 2010, 2015
Bolivia: 1994, 2003, 2008	Mali: 1996, 2001, 2006, 2012
Burkina Faso: 1999, 2010	Morocco: 1992, 2003
Burundi: 2010, 2016	Mozambique: 1997, 2003, 2011
Cambodia: 2000, 2005, 2010, 2014	Namibia: 1992, 2000, 2013
Cameroon: 1998, 2004, 2011	Nepal: 1996, 2006, 2016
Central African Republic: 1994	Niger: 1992, 2006, 2012
Chad: 1997, 2004, 2015	Nigeria: 2008, 2013
Congo, Democratic Republic: 2007, 2013	Peru: 1991, 1996, 2000
Congo, Republic: 2005, 2011	Philippines: 1993, 1998
Côte d'Ivoire: 1994, 2012	Rwanda: 2000, 2005, 2010, 2015
Dominican Republic: 2002, 2007	São Tomé & Príncipe: 2008
Ethiopia: 2000, 2005, 2011, 2016	Senegal: 1993, 2005, 2011
Gabon: 2000, 2012	Sierra Leone: 2008, 2013
Ghana: 2007	South Africa: 1998, 2016
Guinea: 1999, 2005, 2012	Swaziland: 2006
Haiti: 2000, 2006, 2017	Tanzania: 1996, 2004, 2010, 2015
Indonesia: 1994, 1997, 2002, 2007, 2012	Togo: 1998, 2014
Jordan: 1997	Zambia: 1996, 2002, 2007, 2013
Kenya: 1998, 2003, 2009, 2014	Zimbabwe: 1994, 1999, 2005, 2010, 2015

Table A.2. Partial Correlations of Sibling and Child Under-5 Mortality, Women Aged 45-49

	Any child death and any sibling death	# child deaths and # sibling deaths
	(1)	(2)
Within-survey correlation	0.077	0.074
Number of observations	131,518	131,518

Note: Partial correlations are computed after conditioning on survey indicators. Sampling weights are rescaled to reflect each survey's contribution to the sample.

Table A.3. Mothers' vs. Daughters' Reports of Any Under-5 Death

	D's report (%)		
M's report	0	1+	N
0	98.0	2.0	59,339
1+	14.4	85.6	35,552

Note: Sample includes coresident 15-19 year olds and their 30-49 year old mothers when both responded to the survey. Mothers and daughters are interviewed separately and privately. Sampling weights are rescaled to reflect each survey's contribution to the sample.

Table A.4. Mothers' vs. Daughters' Reports of Any Under-5 Death

	D's report (%)							
M's report	0	1	2	3	4	5	6+	N
0	98.0	1.6	0.3	0.1	0.0	0.0	0.0	59,339
1	17.3	80.4	1.9	0.2	0.1	0.0	0.0	21,114
2	11.2	6.5	79.9	2.1	0.3	0.0	0.0	9,172
3	8.3	3.1	7.8	77.7	2.8	0.4	0.0	3,761
4	7.7	2.6	3.1	7.6	77.0	1.5	0.4	1,505

Note: Sample includes coresident 15-19 year olds and their 30-49 year old mothers when both responded to the survey. Mothers and daughters are interviewed separately and privately. Sampling weights are rescaled to reflect each survey's contribution to the sample.

Table A.5. Pooled Birth-Level Logit Estimations by Gender

	Female	Male	Female	Male
	(1)	(2)	(3)	(4)
# sibling U5 deaths	1.09	1.09		
	[.0034]	[.0032]		
Sibs ever born	.98	.98		
	[.0016]	[.0015]		
# female sibling U5 death			1.11	1.09
			[.0058]	[.0056]
Female sibs ever born			.99	.98
			[.0022]	[.0021]
# male sibling U5 death			1.08	1.10
			[.0051]	[.0048]
Male sibs ever born			.98	.98
			[.0022]	[.0021]
AME(# sib. deaths)	.010	.010		
AME(# female sib. deaths)			.011	.010
AME(# male sib. deaths)			.009	.011
Observations	1,276,858	1,333,004	1,276,858	1,333,004

Note: The reported estimates are logit odds ratios. Brackets contain standard errors clustered at the survey cluster level. AME refers to the average marginal effect of the indicated measure of sibling death(s). All models include survey indicators. Sampling weights are rescaled to reflect each survey's contribution to the sample.

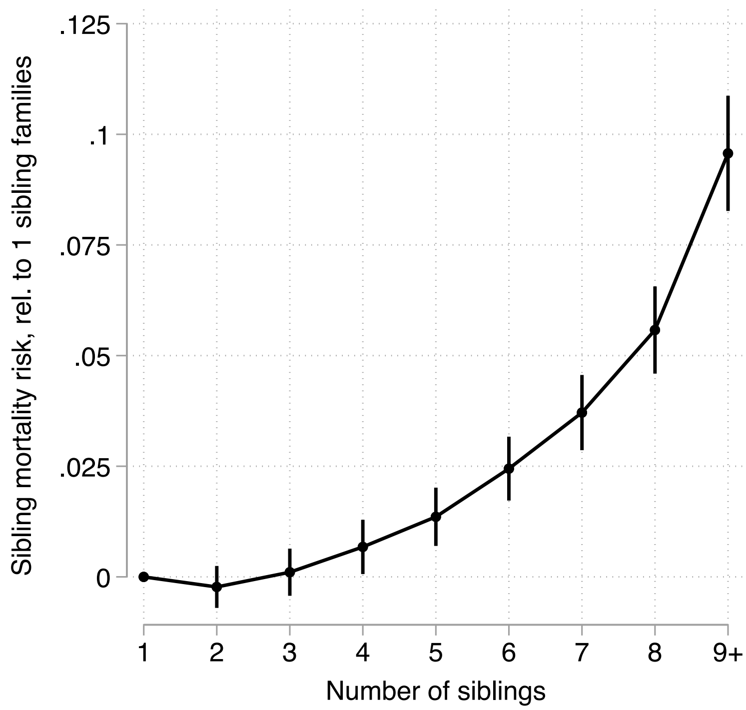


Figure A.1. Sibship Size and Sibling Mortality

Note: This figure plots the relationship between sibship size and sibling mortality rates. The unit of observation is the survey-sibsize cell. We regress the sibling under-5 mortality rate on sibship size indicators and survey indicators. Mortality rates are scaled from 0 to 1. Cells are weighted by the number of women. Spikes are 95% confidence intervals based on standard errors clustered at the country level. Sampling weights are rescaled to reflect each survey's contribution to the cell.

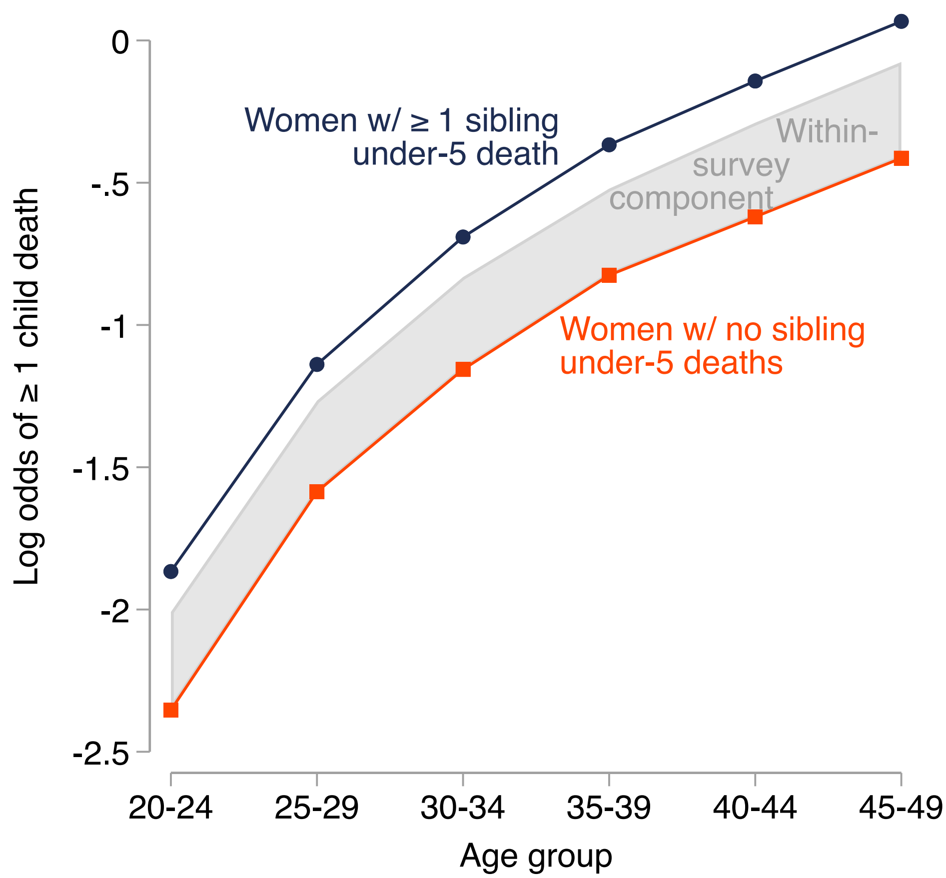


Figure A.2. Log Odds of Any Child Death, by Any Sibling Death

Note: For each five-year age group, we plot the log odds of any under-5 child death separately for women with and without deceased siblings. Sampling weights are rescaled to reflect each survey's contribution to the sample.

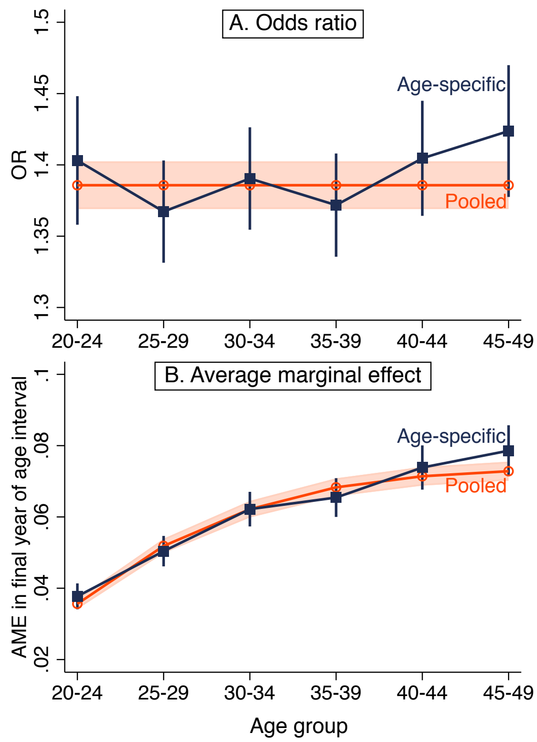


Figure A.3. Mother-Level Logit Results by Age

Note: This figure demonstrates the robustness of the woman-level logit estimates to age-specific estimations. Point estimates and 95% confidence intervals based on women-level logit regressions of any under-5 child death on any under-5 sibling death. All regressions include survey indicators and single-year age indicators. Pooled estimations include women of all ages; age-specific estimations are separate for each five-year age group. Average marginal effects are computed for the final age in each age interval; confidence intervals are based on standard errors computed using the delta method. Sampling weights are rescaled to reflect each survey's contribution to the sample.

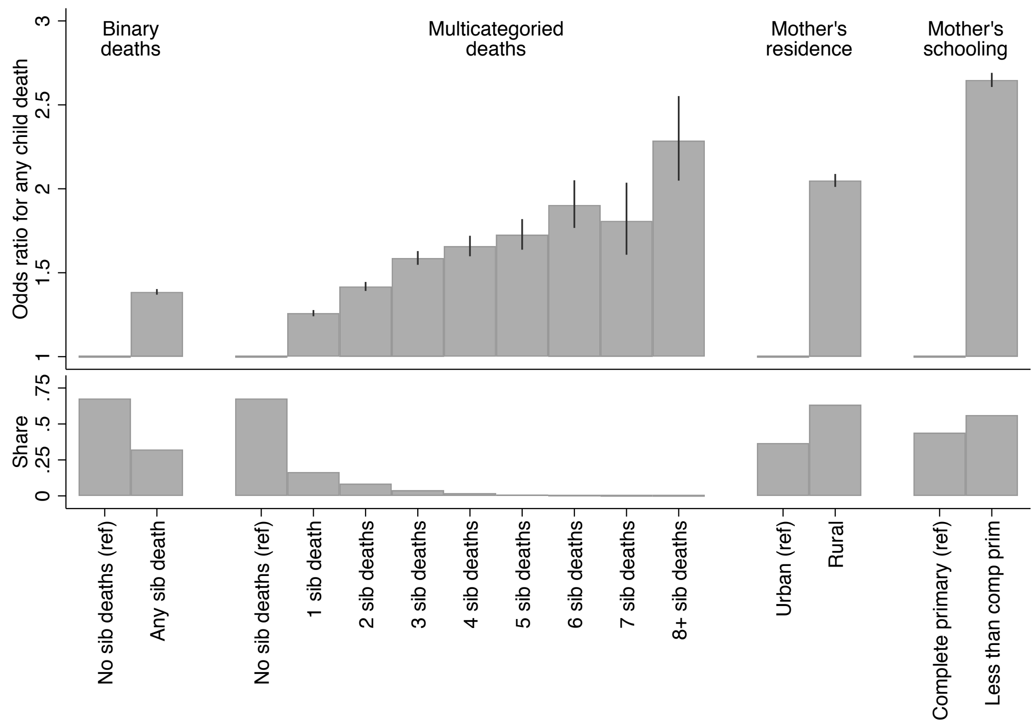


Figure A.4. Comparison with Other Under-5 Mortality Differentials
Note: The top panel presents point estimates and 95% confidence intervals of odds ratios from four woman-level logit regressions of any under-5 child death on the indicated categorical variables in the figure. All regressions include survey indicators and single-year age indicators. The bottom panel presents histograms of the categorical variables. Sampling weights are rescaled to reflect each survey's contribution to the sample.

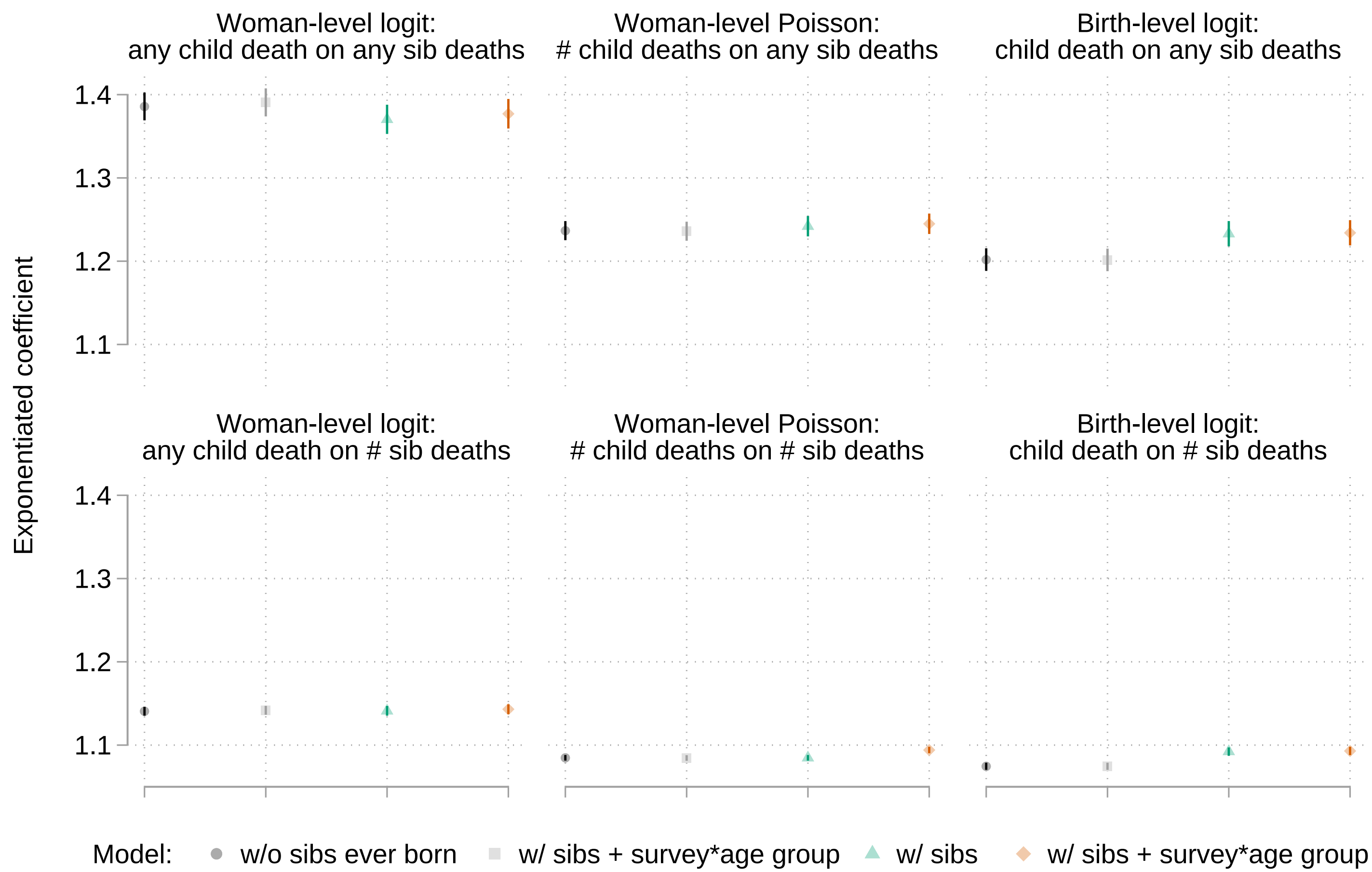


Figure A.5. Robustness to Survey-by-Age Group Effects
Note: We add survey-by-age group indicators to each regression from Table 1.2. We report the new estimates alongside the original estimates from Table 1.2. Spikes represent 95% confidence intervals

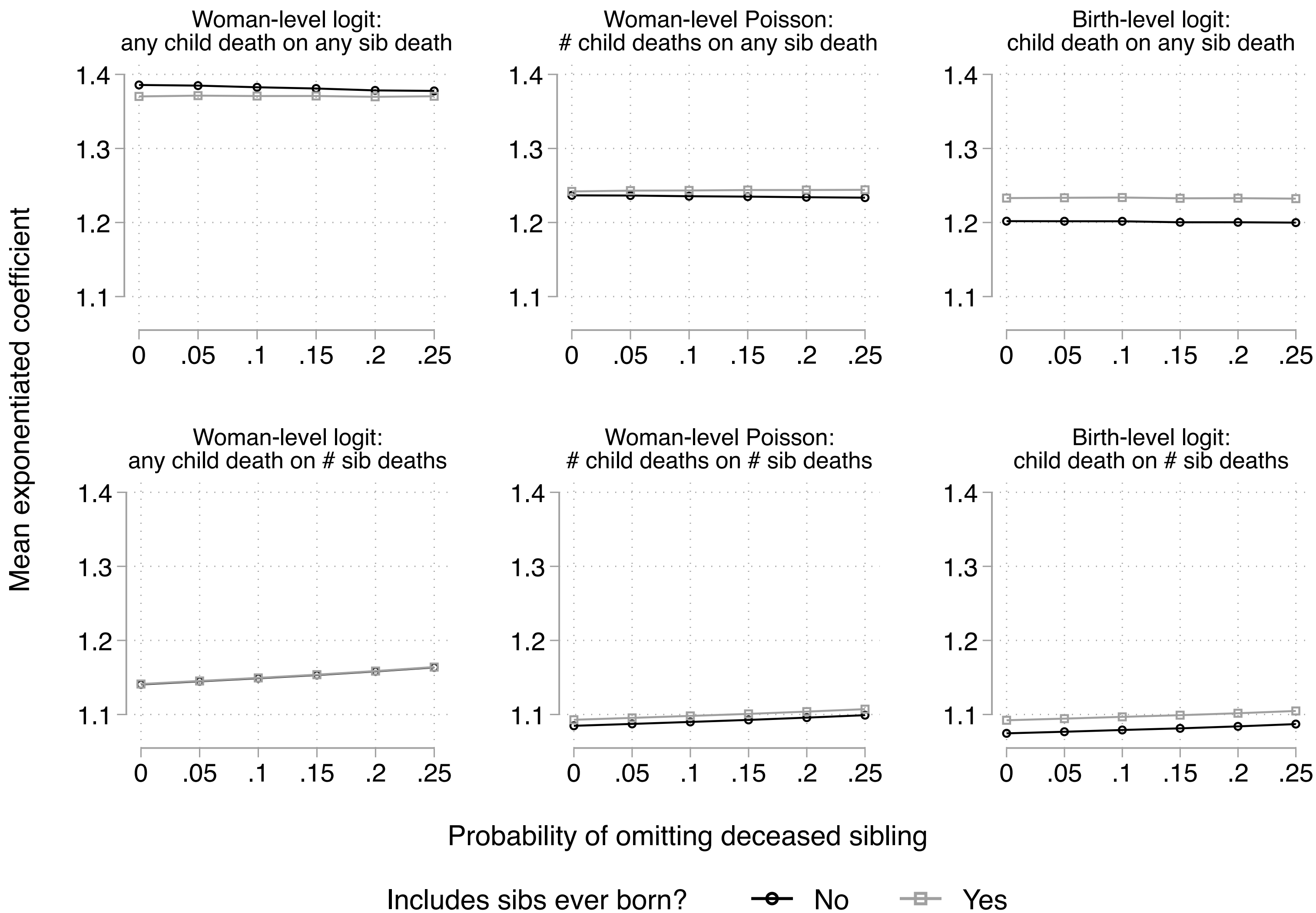


Figure A.6. Monte Carlo Simulations of Measurement Error
Note: This figure reports the impact of simulated measurement error on our mortality persistence estimates. We simulate the omission of reported deceased siblings for different probabilities of omission. For each positive probability, we draw the number of omitted deceased siblings from a binomial distribution 50 times. We estimate each regression from Table 1.2 in each simulated dataset. We plot the mean exponentiated coefficient (odds ratio or incidence rate ratio) across the 50 draws. At p = 0, we plot the result from Table 1.2, with no simulated measurement error.

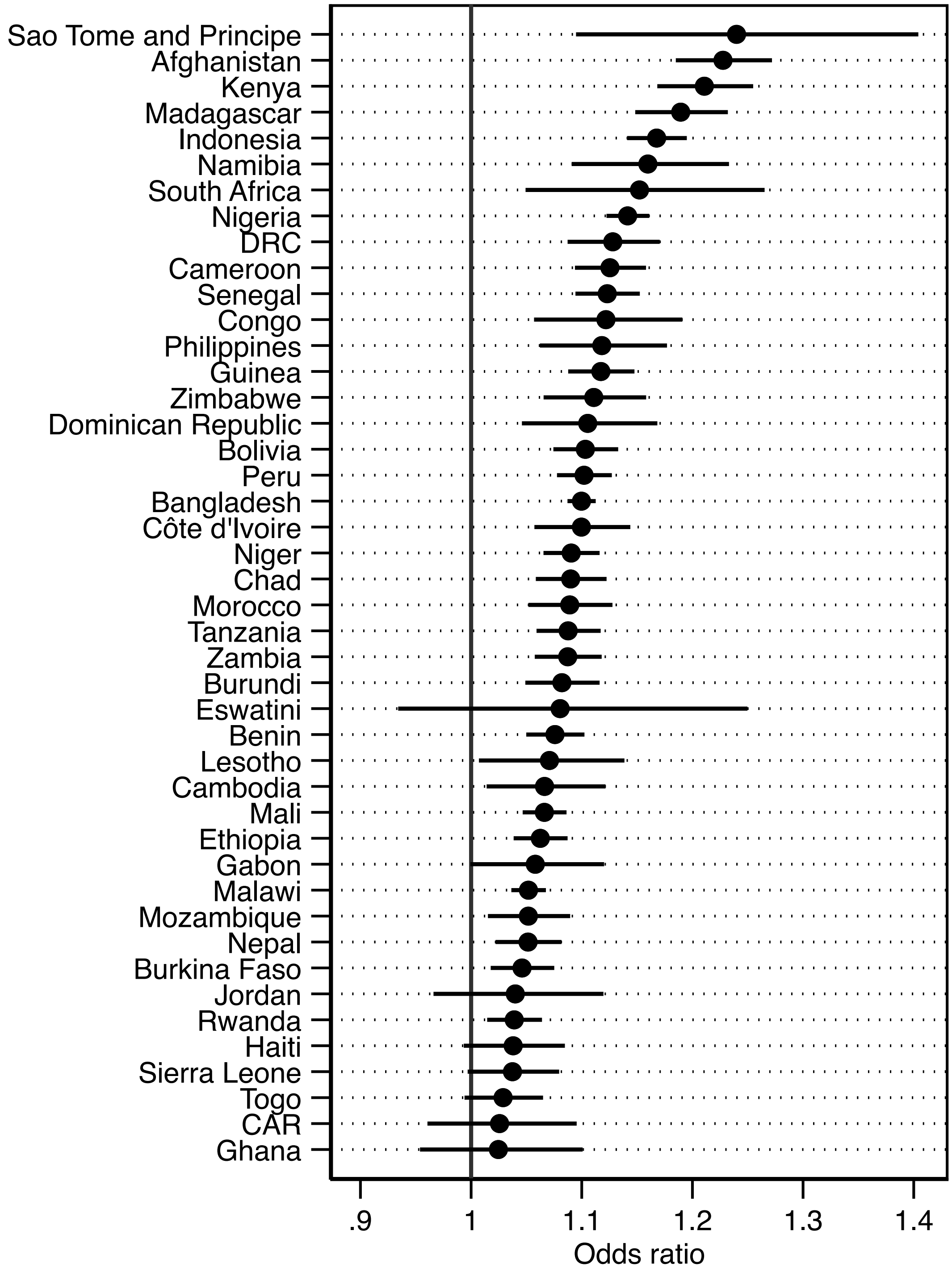


Figure A.7. Mortality Persistence by Country

Note: This figure reports mortality persistence estimates for each country in our sample. The plotted estimates are odds ratios from birth-level logit regressions of under-5 deaths on the mother's number of under-5 sibling deaths. All regressions include the mother's number of siblings ever born and survey indicators. Spikes represent 95% confidence intervals based on standard errors clustered at the survey cluster level. Sampling weights are rescaled to reflect each survey's contribution to each country sample.

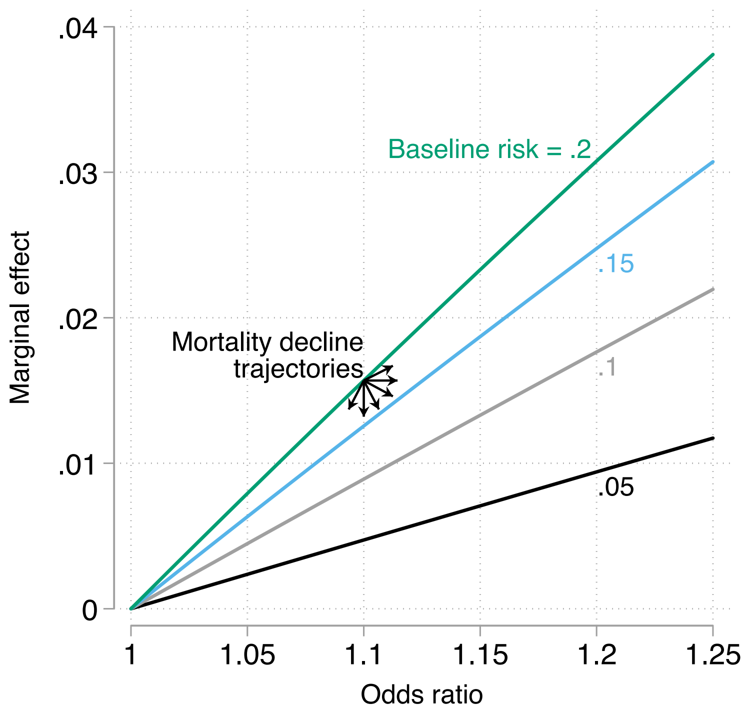


Figure A.8. Absolute Versus Proportional Mortality Persistence for a Binary Risk Factor
Note: Each ray from the origin specifies the relationship between the marginal effect and the odds ratio for a binary risk factor (e.g., any sibling death) at a given level of baseline mortality risk. At higher baseline mortality risk, a given odds ratio translates to a larger marginal effect. The mortality decline trajectories demonstrate possible paths for the odds ratio and marginal effect as mortality falls.

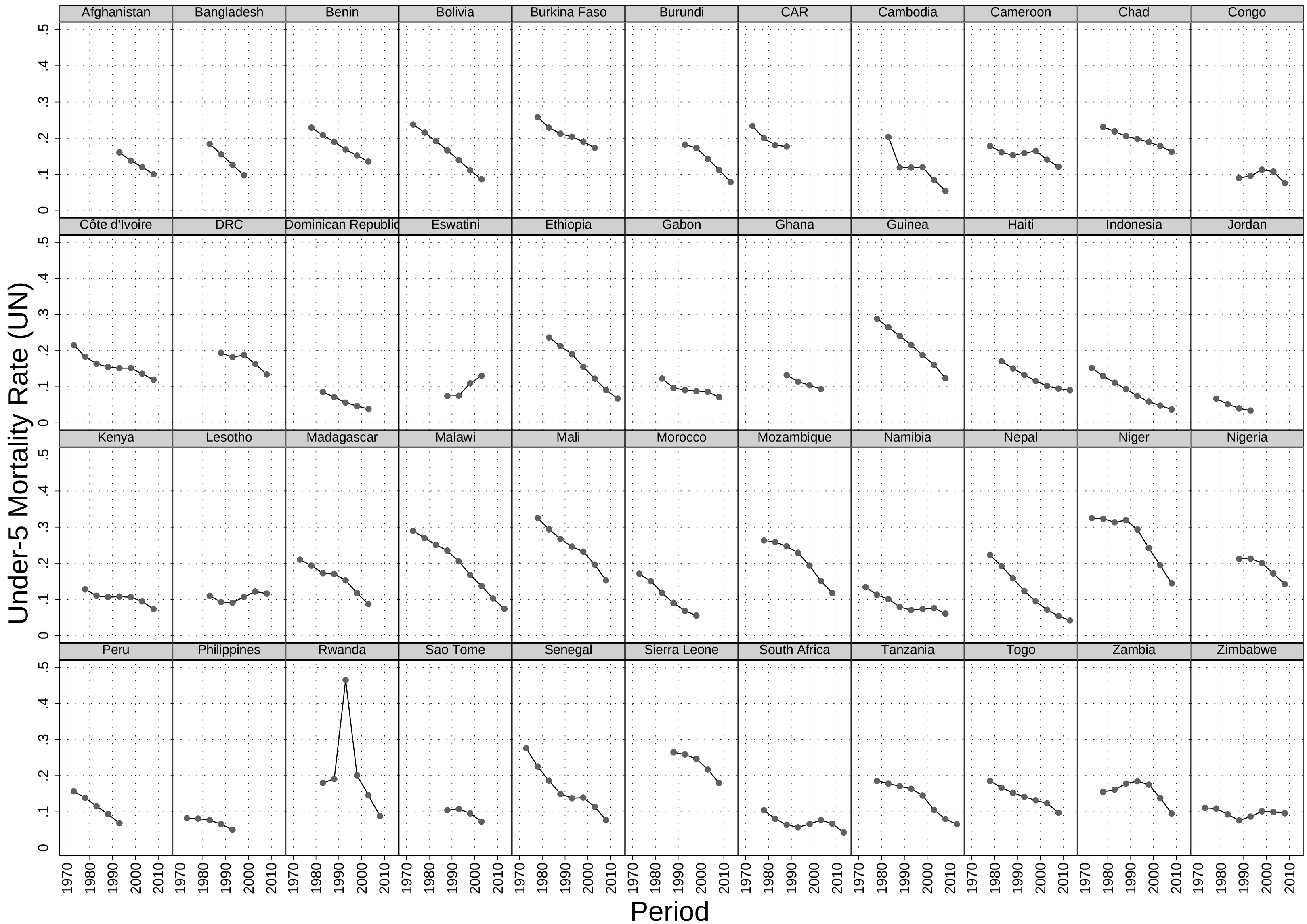


Figure A.9. Under-5 Mortality Rate over Time, by Country
Note: Rates are scaled from 0 to 1.

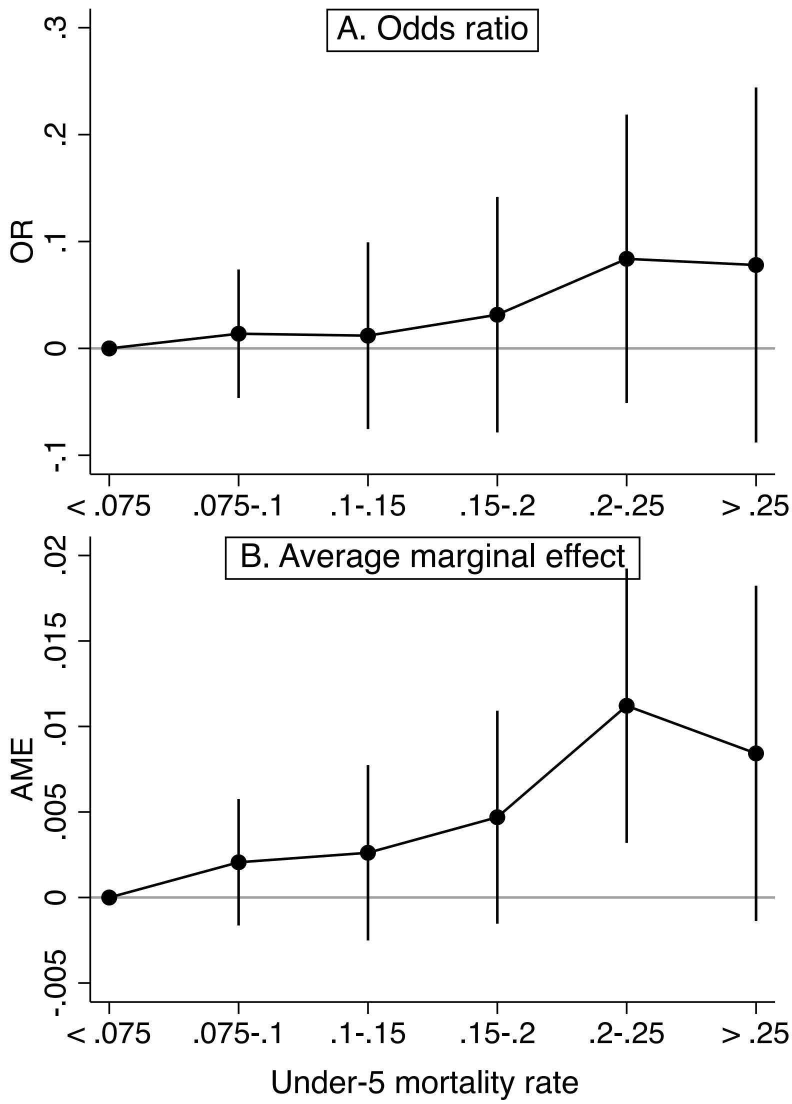


Figure A.10. Semi-Parametric Panel Analyses

Note: The figure replicates Table 1.4, columns (1) and (5), but with under-5 mortality separated into 6 bins. The point estimates are the coefficients for 5 bin indicators, leaving out the lowest as the reference category. Spikes are 95% confidence intervals based on standard errors clustered at the country level. OR is the odds ratio. AME is the average marginal effect. Each panel represents a separate cell-level regression including country and period fixed effects. Panel A corresponds to Table 1.4, column (1), while the Panel B corresponds to Table 1.4, column (3).

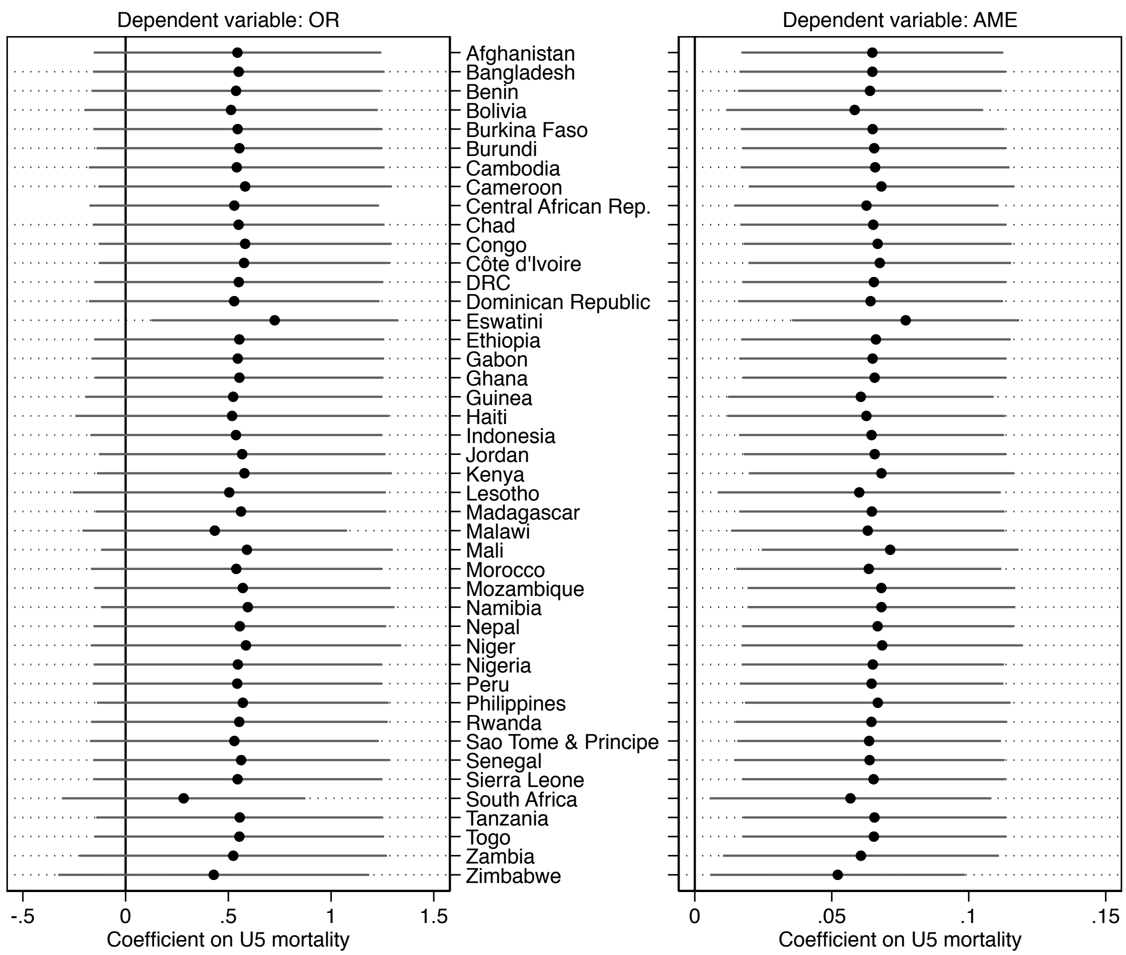


Figure A.11. Leave-One-Out Panel Analyses

Note: This figure replicates Table 1.4, columns (1) and (5), leaving out one country at a time. The point estimates report the cell-level association of the under-5 mortality rate with the intergenerational persistence of under-5 mortality, net of country fixed effects and period fixed effects. Spikes are 95% confidence intervals based on standard errors clustered at the country level. OR is the odds ratio. AME is the average marginal effect. The left-hand panel corresponds to Table 1.4, column (1), while the right-hand panel corresponds to Table 1.4, column (3).

Appendix B for Chapter 2

Table B.1. Pakistan evidence: Age at the time of partition and marriage outcomes (clustered standard errors)

(a) Young is 13 to 17

	Ever married	Child marriage	Marriage age	Married during partition
	(1)	(2)	(3)	(4)
Young X Displaced	.01***	.038	-.28	.053**
	(.002)	(.044)	(.317)	(.021)
Young	-.013***	.06**	-.63***	.23***
	(.002)	(.029)	(.236)	(.016)
Displaced	.00024	.019	-.25	-.012***
	(.001)	(.034)	(.261)	(.004)
Comparison mean	.997	.308	19.3	.0169
Adjusted R^2	.00	.01	.01	.08
N	27406	26654	26654	26654

(b) Young is 1 to 5

	Ever married	Child marriage	Marriage age	Married during partition
	(1)	(2)	(3)	(4)
Young X Displaced	-.014*	-.0043	.14	.012***
	(.008)	(.043)	(.305)	(.004)
Young	-.042***	.13***	-1.2***	-.017***
	(.005)	(.026)	(.214)	(.003)
Displaced	.00024	.019	-.25	-.012***
	(.001)	(.034)	(.261)	(.004)
Comparison mean	.997	.308	19.3	.0169
Adjusted R^2	.01	.01	.02	.01
N	35209	32967	32967	32967

Note: Each column represents a single regression estimated from Equation 2.2. The analysis sample is restricted to displaced (treatment group) and native-born (comparison group) females living in Punjab at the time of enumeration that were ages (a) 13 to 17 or 30 to 32 at the time of partition or (b) 1 to 5 or 30 to 32 at the time of partition. Standard errors clustered at the interaction of whether young at partition, whether displaced, district of residence, and whether the respondent lives in an urban area are reported in parentheses. Individual observations are weighted by inverse sampling probability. The data are from the 1973 Pakistan HED survey. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table B.2. Pakistan evidence: Age at the time of partition and other other own outcomes (clustered standard errors)

(a) Young is 13 to 17

	Literate	Completed primary school before partition	Continued education after partition	Children born	Children surviving
	(1)	(2)	(3)	(4)	(5)
Young X Displaced	.013	.011	-.0027	.43**	.55***
	(.029)	(.017)	(.003)	(.207)	(.141)
Young	.027**	.015*	.0093***	-.42***	-.22**
	(.013)	(.008)	(.002)	(.140)	(.091)
Displaced	.00026	-.0041	-3.6e-15	.34**	-.094
	(.012)	(.007)	(.000)	(.149)	(.103)
Comparison mean	.0284	.0205	0	5.64	4.77
Adjusted R^2	.00	.00	.00	.01	.01
N	27406	26881	26953	27109	25657

(b) Young is 1 to 5

	Literate	Completed primary school	Children born	Children surviving
	(1)	(2)	(3)	(4)
Young X Displaced	.035	.031	-.052	.38***
	(.045)	(.031)	(.177)	(.125)
Young	.064***	.041***	-2.4***	-1.8***
	(.020)	(.014)	(.122)	(.084)
Displaced	.00026	-.0041	.34**	-.094
	(.012)	(.007)	(.149)	(.103)
Comparison mean	.0284	.0205	5.64	4.77
Adjusted R^2	.01	.01	.17	.15
N	35209	34110	33769	30678

Note: Each column represents a single regression estimated from Equation 2.2.The analysis sample is restricted to displaced (treatment group) and native-born (comparison group) females living in Punjab at the time of enumeration that were ages (a) 13 to 17 or 30 to 32 at the time of partition or (b) 1 to 5 or 30 to 32 at the time of partition. Standard errors clustered at the interaction of whether young at partition, whether displaced, district of residence, and whether the respondent lives in an urban area are reported in parentheses. Individual observations are weighted by inverse sampling probability. The data are from the 1973 Pakistan HED survey. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table B.3. Pakistan evidence: Age at the time of partition and spousal characteristics (clustered standard errors)

(a) Young is 13 to 17

	Age difference	Same nativity	Literate	Completed primary school	Continued education after partition	
	(1)	(2)	(3)	(4)	(5)	(6)
Young X Displaced	-.13	-.0032	.0026	.0037	.00085	.025
	(.322)	(.009)	(.055)	(.046)	(.006)	(.035)
Young	.51**	-.011	.086***	.064***	.013***	.14***
	(.222)	(.007)	(.030)	(.023)	(.004)	(.020)
Displaced	.024	-.01**	.019	.012	-.0012**	.0031
	(.271)	(.005)	(.028)	(.023)	(.001)	(.032)
Comparison	6.00	.99	.15	.10	.00	.63
Adjusted R^2	.0011	.0021	.0086	.0065	.0032	.021
N	21807	21804	21807	20429	20950	27406

(b) Young is 1 to 5

	Age difference	Same nativity	Literate	Completed primary school	Continued education after partition
	(1)	(2)	(3)	(4)	(5)
Young X Displaced	-.58*	.011	.054	.057	.065*
	(.340)	(.013)	(.055)	(.050)	(.039)
Young	-.094	-.048***	.17***	.14***	.17***
	(.218)	(.012)	(.032)	(.027)	(.021)
Displaced	.024	-.01**	.019	.012	-.0012**
	(.271)	(.005)	(.028)	(.023)	(.001)
Comparison	6.00	.99	.15	.10	.00
Adjusted R^2	.00083	.007	.027	.024	.051
N	27390	27387	27390	25226	24930

Note: Each column represents a single regression estimated from Equation 2.2. The analysis sample is restricted to displaced (treatment group) and native-born (comparison group) females living in Punjab at the time of enumeration that were ages (a) 13 to 17 or 30 to 32 at the time of partition or (b) 1 to 5 or 30 to 32 at the time of partition. The dependent variable "Age difference" is the difference between the spouse's age and the respondent's age. The dependent variable "Same nativity" is an indicator of whether the respondent's spouse has the same nativity (born in India or Pakistan) as the respondent. The dependent variable "Literate" is an indicator of whether the respondent's spouse is literate (regardless of formal school attendance). The dependent variable "Completed primary school" is an indicator of whether the respondent's spouse completed primary school. The dependent variable "Continued education after partition" is an indicator of whether the spouse's highest level of education attainment occurred after partition. Standard errors clustered at the interaction of whether young at partition, whether displaced, district of residence, and whether the respondent lives in an urban area are reported in parentheses. Individual observations are weighted by inverse sampling probability. The data are from the 1973 Pakistan HED survey. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table B.4. Pakistan evidence: Robustness of main results to alternative choices of "young" group

(a) Marriage outcomes

	Ever married	Child marriage	Marriage age	Married during partition
	(1)	(2)	(3)	(4)
Young is 13 to 17 (reference)	.01***	.038**	-.28**	.053***
	(.002)	(.016)	(.126)	(.009)
Young is 1 to 5 (reference)	-.014***	-.0043	.14	.012***
	(.004)	(.016)	(.124)	(.003)
Young is 1 to 9	-.0015	.0084	.043	.012***
	(.003)	(.015)	(.119)	(.003)
Young is 6 to 12	.016***	.027	-.12	.014***
	(.003)	(.016)	(.125)	(.003)
Young is 10 to 17	.011***	.035**	-.22*	.046***
	(.002)	(.015)	(.120)	(.006)
Young is 10 to 19	.012***	.034**	-.22*	.05***
	(.002)	(.015)	(.120)	(.006)
Comparison mean	.997	.308	19.3	.0169

(b) Other own outcomes

	Literate	Completed primary school before partition	Continued education after partition	Children born	Children surviving
	(1)	(2)	(3)	(4)	(5)
Young is 13 to 17 (reference)	.013**	.011**	-.0027**	.43***	.55***
	(.005)	(.004)	(.001)	(.101)	(.078)
Young is 1 to 5 (reference)	.035***	.031***	.027***	-.052	.38***
	(.006)	(.005)	(.004)	(.095)	(.074)
Young is 1 to 9	.02***	.018***	.013***	.15	.5***
	(.005)	(.004)	(.003)	(.092)	(.072)
Young is 6 to 12	.0062	.0071	.0011	.26***	.55***
	(.006)	(.004)	(.003)	(.098)	(.077)
Young is 10 to 17	.012**	.01***	.000011	.38***	.56***
	(.005)	(.004)	(.001)	(.095)	(.074)
Young is 10 to 19	.012**	.01***	-.000064	.39***	.57***
	(.005)	(.004)	(.001)	(.095)	(.073)
Comparison mean	.0284	.0205	0	5.64	4.77

Note: Each cell reports the difference-in-difference coefficient δ estimated from Equation 2.2. Our analysis sample is restricted to displaced (treatment group) and native-born (comparison group) females living in Punjab at the time of enumeration that were ages in the indicated "young" age range (see row name) or 30 to 32 at the time of partition Panel (a) reports the robustness of our main marriage outcomes (same outcomes as Table 2.2). Panel (b) reports the robustness of the women's other own outcomes (same outcomes as Table 2.3a).The dependent variable "Ever married" is an indicator of whether the respondent has ever been married. The dependent variable "Child marriage" is an indicator whether the respondent was married before the age of 18. The dependent variable "Marriage age" is the respondent's age at first marriage. The dependent variable "Married during partition" is an indicator variable of whether a respondent was first married between 1946 - 1949. The dependent variable "Literate" is an indicator of whether the respondent is literate (regardless of formal school attendance). The dependent variable "Completed primary school (before partition)" is an indicator of whether the respondent completed primary school (before partition). The dependent variable "Children ever born" is the total number of children ever born to the respondent. The dependent variable "Children surviving " is the number of surviving children the respondent has at the time of the survey. Heteroskedasticity-robust standard errors are reported in parentheses. Individual observations are weighted by inverse sampling probability. The data are from the 1973 Pakistan HED survey. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table B.5. Pakistan evidence: Robustness of main results to alternative choices of "old" comparison group

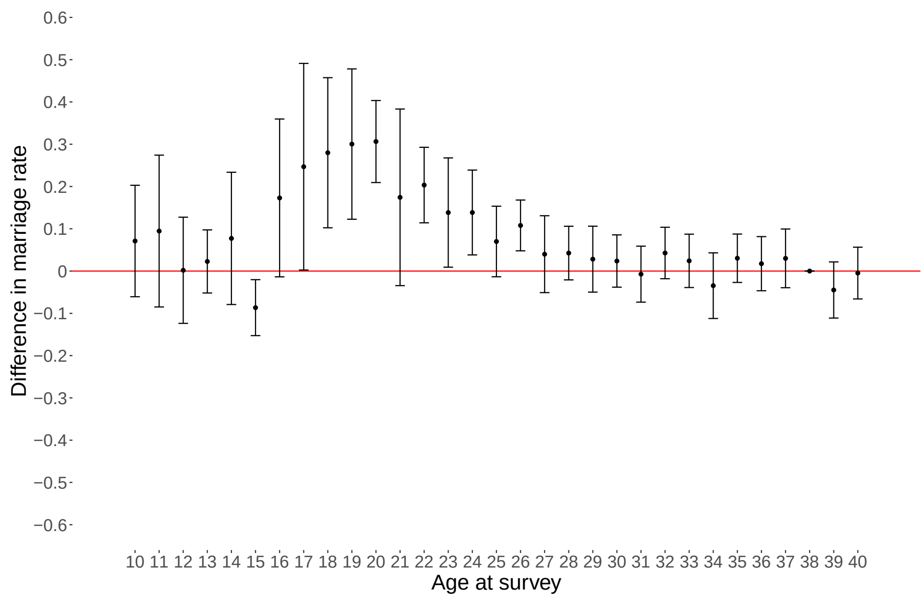
(a) Marriage outcomes

	Ever married	Child marriage	Marriage age	Married during partition
	(1)	(2)	(3)	(4)
Old is 30 to 32 (reference)	.01***	.038**	-.28**	.053***
	(.002)	(.016)	(.126)	(.009)
Old is 30 to 35	.0046**	.036***	-.21**	.048***
	(.002)	(.013)	(.101)	(.008)
Old is 30 to 40	-.0004	.044***	-.23**	.046***
	(.002)	(.012)	(.094)	(.008)
Old is 30 to 50	-.0011	.046***	-.22**	.045***
	(.002)	(.011)	(.091)	(.008)
none				

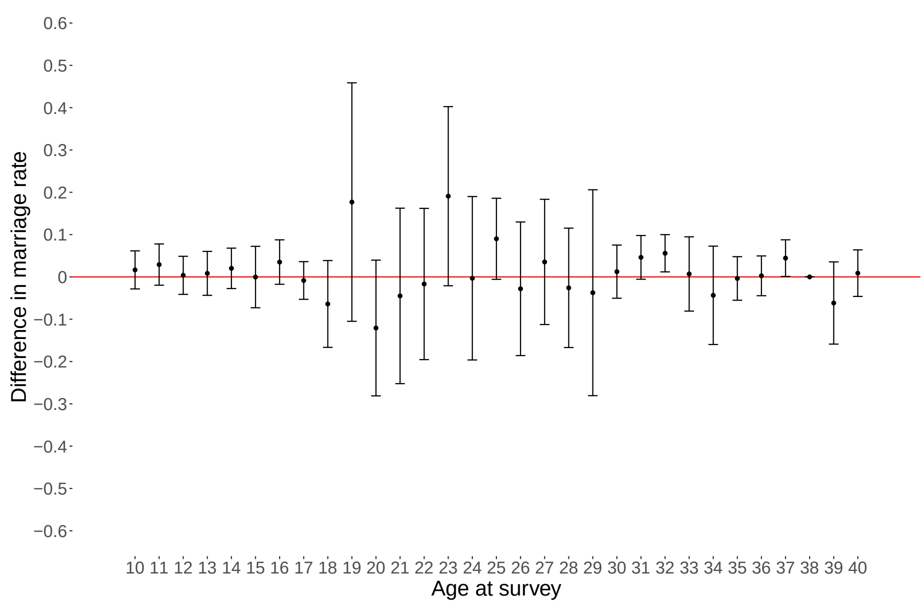
(b) Other own outcomes

	Literate	Completed primary school before partition	Continued education after partition	Children born	
	(1)	(2)	(3)	(4)	(5)
Old is 30 to 32 (reference)	.013**	.011**	-.0027**	.43***	.55***
	(.005)	(.004)	(.001)	(.101)	(.078)
Old is 30 to 35	.014***	.012***	-.0027**	.47***	.57***
	(.004)	(.004)	(.001)	(.078)	(.060)
Old is 30 to 40	.027***	.022***	-.0027**	.46***	.58***
	(.004)	(.004)	(.001)	(.073)	(.056)
Old is 30 to 50	.03***	.022***	-.0027**	.47***	.57***
	(.004)	(.003)	(.001)	(.068)	(.053)
none					

Note: Each cell reports the difference-in-difference coefficient δ estimated from Equation 2.2. Our analysis sample is restricted to displaced (treatment group) and native-born (comparison group) females living in Punjab at the time of enumeration that were ages 13 to 17 or in the indicated comparison age range (see row name) at the time of partition. Panel (a) reports the robustness of our main marriage outcomes (same outcomes as Table 2.2). Panel (b) reports the robustness of the women's other own outcomes (same outcomes as Table 2.3a).The dependent variable "Ever married" is an indicator of whether the respondent has ever been married. The dependent variable "Child marriage" is an indicator whether the respondent was married before the age of 18. The dependent variable "Marriage age" is the respondent's age at first marriage. The dependent variable "Married during partition" is an indicator variable of whether a respondent was first married between 1946 - 1949. The dependent variable "Literate" is an indicator of whether the respondent is literate (regardless of formal school attendance). The dependent variable "Completed primary school (before partition)" is an indicator of whether the respondent completed primary school (before partition). The dependent variable "Children ever born" is the total number of children ever born to the respondent. The dependent variable "Children surviving " is the number of surviving children the respondent has at the time of the survey. Heteroskedasticity-robust standard errors are reported in parentheses. Individual observations are weighted by inverse sampling probability. The data are from the 1973 Pakistan HED survey. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$



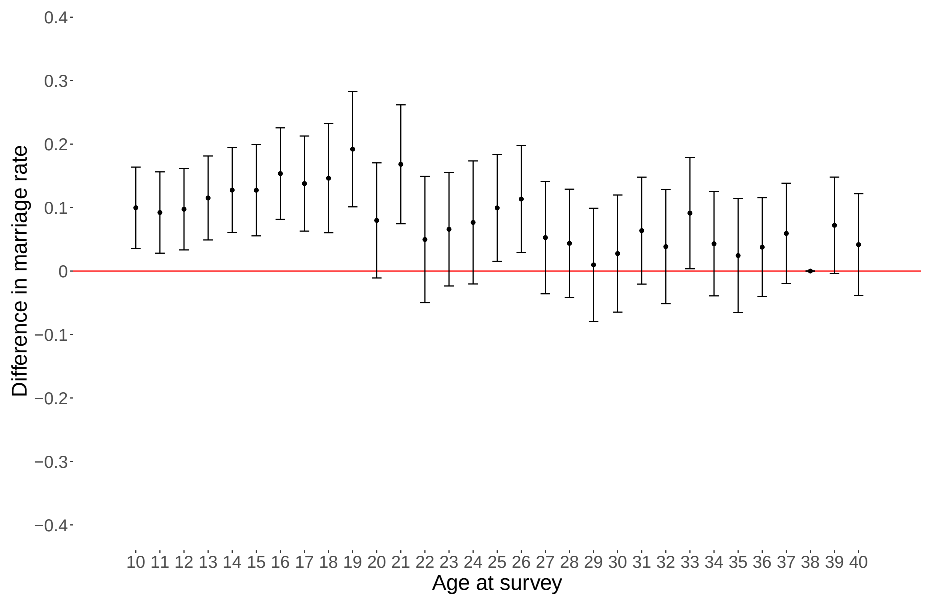
(a) Female



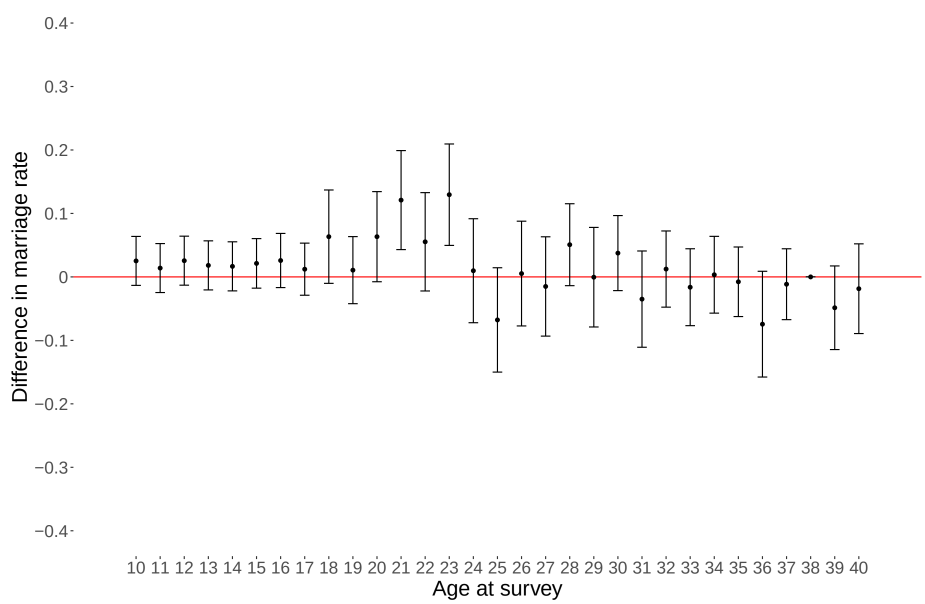
(b) Male

Figure B.1. Cross-country evidence: Effect of displacement on marriage rates (destination fixed effects)

Note: This figure reports differences in age-specific marriage rates between recently displaced (displaced within 5 years of survey) and non-migrant individuals of a given gender as indicated in each panel. The dependent variable is whether the respondent has ever been married. Each dot represents the coefficients δ_a estimated from Equation 2.1 with destination X country X year fixed effects. We define "destination" to be the largest geographical sub-unit reported in each datasets for the residence of the respondent at the time of the survey.The error bars report a 95% for each coefficient estimate where standard errors are heteroskedasticity-robust standard errors. Individual observations are weighted by inverse sampling probability.



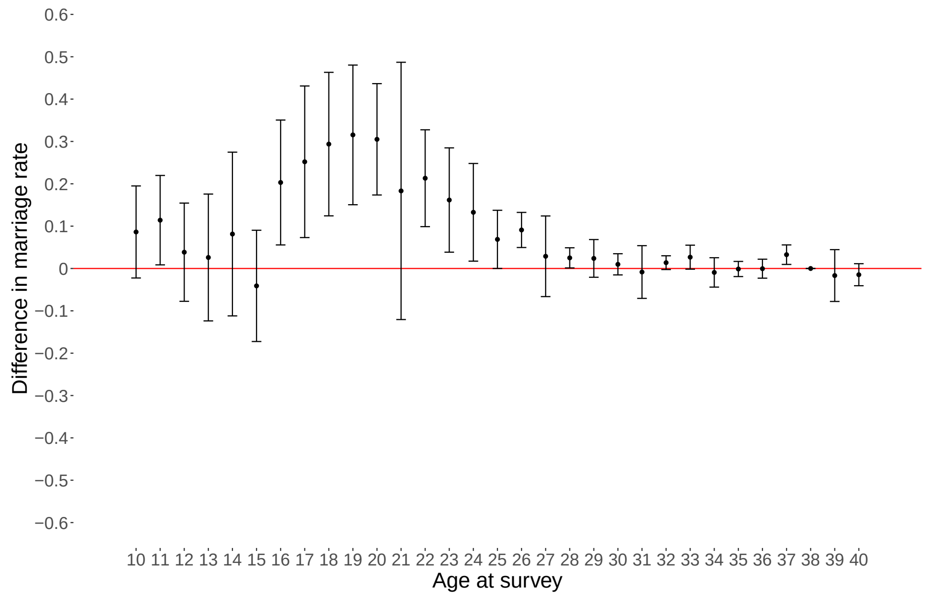
(a) Female



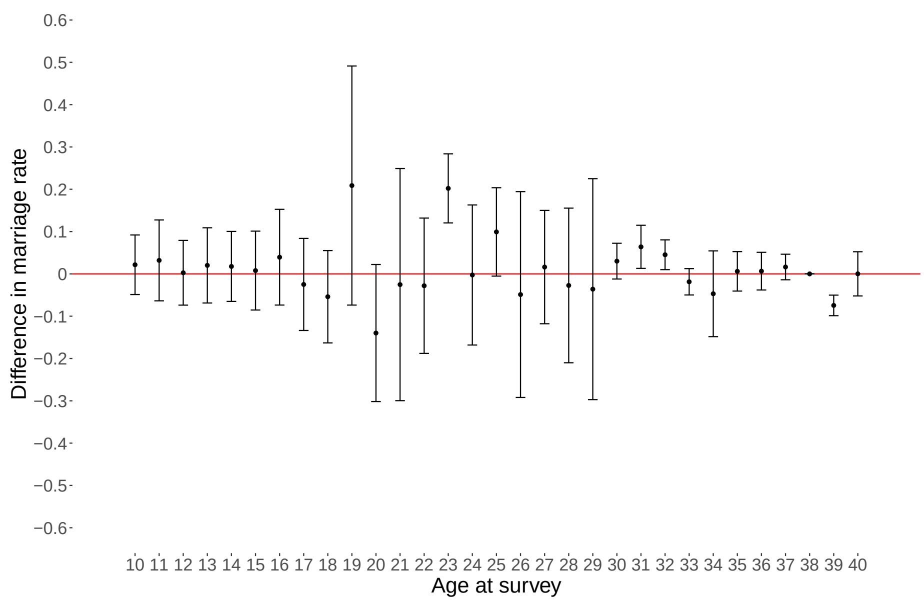
(b) Male

Figure B.2. Cross-country evidence: Effect of displacement on marriage rates (birthplace fixed effects)

Note: This figure reports differences in age-specific marriage rates between recently displaced (displaced within 5 years of survey) and non-migrant individuals of a given gender as indicated in each panel. The dependent variable is whether the respondent has ever been married. Each dot represents the coefficients δ_a estimated from Equation 2.1 with birthplace X country X year fixed effects. We define "birthplace" to be the largest geographical sub-unit reported in the dataset for the birthplace of the respondent. The error bars report a 95% for each coefficient estimate where standard errors are heteroskedasticity-robust standard errors. Individual observations are weighted by inverse sampling probability.



(a) Female



(b) Male

Figure B.3. Cross-country evidence: Effect of displacement on marriage rates (clustered standard errors)

Note: This figure reports differences in age-specific marriage rates between recently displaced (displaced within 5 years of survey) and non-migrant individuals of a given gender as indicated in each panel. The dependent variable is whether the respondent has ever been married. Each dot represents the coefficients δ_a estimated from Equation 2.1 with country X year fixed effects. The error bars report a 95% for each coefficient estimate where standard errors are clustered at the country X survey Year X displaced level. Individual observations are weighted by inverse sampling probability.

Appendix C for Chapter 3

Table C.1. Survey list

Country	Survey year
Benin	2017
Burkina Faso	2010*
Cameroon	2004, 2011, 2018*
Côte d'Ivoire	2012*
Democratic Republic of the Congo	2013
Ethiopia	2008*
Ghana	2008*
Kenya	2009*, 2014*
Malawi	2004*, 2010*, 2015*
Mali	2006*, 2012*, 2018*
Mozambique	2011
Nigeria	2008*, 2013*, 2018*
Sierra Leone	2013, 2019*
Togo	2014*
Uganda	2011, 2016
Zambia	2013*, 2018*

Table C.2. Pooled estimates: matched partner sample

	X violence with current partner?			
	(1) Any	(2) Physical	(3) Sexual	(4) Emotional
Parent IPV	.239***	.212***	.099***	.154***
	[.0076]	[.007]	[.0057]	[.0071]
Comparison Mean	.29	.19	.07	.20
Observations	48953	48953	48953	48953
R^2	.04	.04	.02	.02

Note: Each column represents a single regression estimated from Equation 3.1. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table C.3. Lee Bounds for selection out of partnership

	X violence with current partner?			
	(1) Any	(2) Physical	(3) Sexual	(4) Emotional
Lower bound	.257***	.234***	.111***	.171***
	[.0032]	[.0032]	[.0024]	[.0031]
Upper bound	.258***	.235***	.112***	.172***
	[.0032]	[.0032]	[.0025]	[.0031]
Comparison Mean	.30	.20	.07	.21
Observations	136438	136438	136438	136438