UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Essays in Development Economics

Permalink
https://escholarship.org/uc/item/07k5h1zg

Author
Siddiqui, Sameem

Publication Date
2020

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Development Economics

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

in

Economics

by

Sameem Iqbal Siddiqui

Committee in charge:

Professor Eli Berman, Co-Chair
Professor Karthik Muralidharan, Co-Chair
Professor Michael J. Callen
Professor Craig McIntosh
Professor Charles D. Sprenger

2020

Copyright
Sameem Iqbal Siddiqui, 2020
All rights reserved.

The dissertation of Sameem Iqbal Siddiqui is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Co-Chair

University of California San Diego

2020

DEDICATION

To my mother Tayyaba and my wife Fatima, the people in my life whose never ending strength and support have brought me where I am today.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to first acknowledge my committee co-chairs Professors Karthik Muralidharan and Eli Berman who have provided overarching support and guidance throughout my graduate degree. This acknowledgment extends to my other committee members, Professors Micheal J. Callen, Craig McIntosh and Charles D. Sprenger, each of whom have lent their expertise on key aspects of my dissertation. Learning and interacting with them has been a truly enlightening experience which I will always cherish.

Chapter 1 uses data which I gained access to through Teevrat Garg, Gordon McGord and Maulik Jagnani to whom I'm very grateful for.

Chapter 2 is coauthored with Prashant Bharadwaj and Frances Lu and has been prepared for submission for publication. I am grateful to Frances for her outstanding contribution to this chapter, and to Prashant for his guidance and support.

Chapter 3 uses survey data collected by Ali Cheema, Asim I. Khwaja, Jacob N. Shapiro and Farooq Naseer through the Center for Economic Research in Pakistan, which they generously gave me access to.

I am also very grateful to all my professors at UCSD for their support and generosity. Last, but not least, I am grateful for my friends and colleagues at UCSD whose company and camaraderie was a great source of support during this time.

VITA

2006-2009	B.Sc. in Mathematics & Economics, University of London International Programmes.
2009-2011	M.Sc. in Economics, Lahore University of Management Sciences, Lahore, Pakistan
2011-2015	Research Associate, Center for Economic Research in Pakistan, Lahore, Pakistan
2015-2020	Teaching Assistant, University of California San Diego, San Diego, U.S.A.
2015-2020	Ph.D. in Economics, University of California San Diego, San Diego, U.S.A.

ABSTRACT OF THE DISSERTATION

Essays in Development Economics

by

Sameem Iqbal Siddiqui

Doctor of Philosophy in Economics

University of California San Diego, 2020

Professor Eli Berman, Co-Chair
Professor Karthik Muralidharan, Co-Chair

This dissertation is composed of three unrelated chapters. Chapter 1 estimates the causal effect of the increase and decrease in ambient light relative to the clock on homicides in Mexico. It does so by using the discontinuous nature of the start and end of daylight saving time (DST) in a regression discontinuity design. It finds no effect on homicides when DST starts in Spring, but finds that transitions out of DST in Fall cause homicides to increase by nearly 7.8% which equates to almost 4 homicides per year with a yearly social cost of 145.2 million Mexican *pesos*. Neither Spring, nor Fall DST transitions influence homicides involving organized crime.

To learn more about the long-term consequences of displacement on women, Chapter

2 examines the marriage market outcomes of forcibly displaced women. It uses data from representative surveys in 7 countries to document that women who are adolescents at the time of displacement are more likely to be married with no such pattern for displaced adolescent men. It provides more robust evidence by using the specific instance of displacement due to the partition of India in 1947. It employs difference-in-differences to find that women who were adolescents when they were displaced by partition were significantly more likely to marry earlier (in line with the cross country evidence), were less likely to continue their education, had more children overall, and do not appear to have married spouses with worse observable characteristics.

Finally, chapter 3 tries to estimate the effect of a large scale anti-poverty cash transfer program in Pakistan, *Benazir Income Support Program* (BISP), on voter turnout and electoral choice. Program eligibility is determined by a threshold on a proxy means test (PMT) score. However, manipulation of the PMT score around the threshold makes the use of traditional regression discontinuity design invalid. In the absence of conclusive causal evidence through bounding techniques, the chapter provides biased point estimates. These show that the correlation between program eligibility and voter turnout, and the correlation between program eligibility and the likelihood for voting for the incumbent party in the 2013 general elections is about 10% point.

Chapter 1

Sunset Times and Homicides

1.1 Introduction

Ambient light can have an impact on human behavior in a number of ways, the most obvious of which is sleep and attentiveness during the day. For example, Jagnani (2018) shows that later sunset times due to geography and seasonality reduces children's sleep which hampers educational attainment. Another example is given by Carrell et al. (2011) who show that school start and end times have an impact on students' academic performance.

Policies that dictate clock settings try to synchronize the timing of social norms (waking up, school timings, work timings, etc.) with ambient light. One such policy is Daylight Saving Times (DST) which is the practice of advancing clocks during warmer months so that darkness falls later in the day according to the clock. While initially adopted for the economic benefit associate with energy conservation, DST imposes some form of social cost in areas and countries where it is practiced. Indeed, Smith (2016) shows that from 2002-2011 the transition into DST during Spring caused over 30 deaths in the United States (US) at a social cost of $275 million annually.

Another potential imposition DST could have on society is through its effect on crime, or how criminals respond to it. Ambient light could be a deterrent towards criminal activity by making it easier to witness crime and, in turn, increasing the likelihood of capture. Alternatively, more light, especially during evenings, could potentially serve as a conduit for criminal activity. Increased outdoor foot traffic in day lit evenings means higher supply of potential victims and reduced search costs for criminals.

Doleac and Sanders (2015) use a regression discontinuity (RD) design to show that transition to DST in Spring (when the sun sets an hour later) causes a 7% and 11% drop in instances of robberies and rape within their analysis sample in the US. They find no effects on aggravated assaults and murder, and no effects of transitions out of DST in Fall (when the sun sets an hour earlier) on any of these crimes. Their results establish that ambient light has statistically and economically significant effects on criminal behavior in a high income setting with effective policing and deterrence. How would crime be effected by ambient light in a middle income and high-crime setting where policing and deterrence is relatively ineffective? I use an RD design to test the impact of increasing and decreasing ambient light brought on by transitions into DST in Spring and out of DST in Fall respectively on homicides in Mexico.

Mexico provides a unique context for this study because of its history of violence associated with organized drug trafficking; ineffective deterrence as indicated by (relatively) low arrest rates; Enforcement of DST in most of the country during the study period (1998-2012); and availability of granular administrative data on the location and timing of homicides and drug trade organization (DTO) related homicides.

RD results for Spring transitions into DST show no effect on later sunset times on homicides. This is in line with Doleac and Sanders (2015) and their (null) results for murder. However, contrary to Doleac and Sanders (2015), my RD results for Fall transitions out of DST show that earlier sunset times cause nearly 7.8% higher homicides, a majority of which occur in urban municipalities. This equates to almost 4 additional homicides per year on average. Neither Spring, nor Fall transitions have any effect on DTO related homicides. These findings are robust to various bandwidths and specifications. Null results for placebo tests that assign “false” DST transitions to dates around the “true” DST transition date provide further confidence in my main statistically significant result. Additional analysis provides suggestive evidence that results for Fall transitions in urban municipalities are driven by areas where, prior to the DST transition, the sun sets during commuting hours. Furthermore, using the social cost of a homicide, I estimate that Fall DST transitions have a yearly social cost of 145.2 million Mexican *pesos* (US $10.8 million) in 2013 Mexican *pesos*.

The rest of this chapter is organized as follows: section 1.2 provides background on DST, homicides and policing in Mexico. Section 1.3 provides a framework for how ambient light may influence criminal behavior. Section 1.4 and 1.5 describe the data and empirical strategy respectively. Section 1.6 provides results and robustness checks. Section 1.7 concludes.

1.2 Background

1.2.1 Daylight Saving Time

DST is the practice of advancing clocks during warmer months so that darkness falls later in the day according to the clock. During this time, clocks are set forward by one hour starting in Spring (typically at 2am on the first day of DST). Clocks are then set one hour back in Fall when DST ends (typically at 3am on the day after the last day of DST). Hence, DST effectively causes an exogenous change in the time when the sun sets.

Adoption of DST within and across different countries was intermittent and sparse until the 1970s when it was widely adopted to conserve energy during the 1970s energy crisis. Mexico, however, adopted DST in 1996 in order to decrease energy consumption and to facilitate commerce and tourism with the US.

The decision to implement Daylight Saving Time (DST) in Mexico now falls under State governments as opposed to a federally established standard in the United States (US). Among the

32 states, 30 observed DST throughout the study period (1998-2012), 1 state (Sonora) stopped observing DST in 1999, and 1 state (Quintana Roo) stopped observing DST in 2012. Also compared to the US, DST in Mexico normally starts later (approximately 2-3 weeks) and ends earlier (approximately 1 week). However, certain states and municipalities that border the U.S. strayed from the nationwide DST start and end dates to match the US. These include the state of Baja California, which did so for one year in 2001 and then 2010 onwards, and 8 municipalities from 5 other states that are close to the US border.

For this study I only focus on states and municipalities that implement DST as per the nationwide DST start and end dates for the main analysis. I use data from states that do not practice DST for placebo tests. I elaborate on this in subsection 1.6.4 and the Appendix.

1.2.2 Homicides and Policing

Homicides

Violent crime has been one of the most pressing concerns to the public in Mexico, the root cause of which is attributed to the production and trafficking of drugs between Latin America and the United States.

Specifically within the study period (1998-2012), homicide rates during the late 1990s were nearly 14 per 100,000, almost twice those in the US (UNODC, 2013). The early 2000s saw a decrease and reached 8.7 per 100,000 in 2004 relative to 5.5 per 100,000 in the US in the same year (UNODC, 2013). However, the second half of the 2000s came with a rapid increase brought on by President Felipe Calderon's war on drugs which started in late 2006. Homicide rate went as high as 22.9 in 2011 (see figure 1.2) with violent civilian deaths per capita reaching higher than in Iraq and Afghanistan during the same period. During this time, over 85 percent of drug-related violence consisted of people involved in the drug trade (Williams, 2012). This violence varied significantly across the country with some states in 2013 seeing homicide rates as low as 11 and

12 (per 100,000) (Yucatán and Baja California Sur respectively) and some experiencing rates as high as 77 and 79 (Morelos and Guerrero respectively) (IEP, 2013).

Policing

Similar to the US, Mexico has state and federal criminal justice systems. While homicides typically fall under state jurisdiction, drug related and organized crimes are often federal offenses.

The arrest rate for homicides in Mexico is low, which is attributed to difficulties with both investigations and apprehension of suspects. In 2012 on average only 36 percent of warrants resulted in an arrest and overall only 20 percent of homicides resulted in an arrest (Evalúa, 2012). This varies significantly by state. In 2010 the states of Chihuahua, Durango and Sinaloa saw clearance rates of 3.6, 4.6 and 7 percent respectively; whereas Baja California Sur and Yucatán had clearance rates of 90 and 100 percent respectively. In contrast, the clearance rate for homicides in the US is approximately 67 percent, ranging from 52 percent to 89 percent across states (Dell, 2015).

1.3 Ambient Light and Criminal Behavior

I consider a simple framework of criminal behavior where criminals act if the expected benefit of crime is greater than the expected costs (Becker, 1968). Ambient light plays a role through factors affecting the latter.

I assume expected cost of crime as a function that increases in the probability of being caught and the severity of punishment, and decreases in the search cost of potential victims. More light during DST, when the sun sets an hour later, means a higher probability of the crime being witnessed and hence being caught. On the other hand, more light also means that people are likely to stay out longer making it easier to search for potential victims. The latter pathway could be especially potent in a context where apprehension is typically low. Less ambient light outside of DST would have the converse effect. Unfortunately, I am unable to disentangle the two effects and the empirical exercise will try and identify the net effect of an increase and decrease in ambient light from transition into and out of DST respectively.

Working with a similar framework, Doleac and Sanders (2015) find that transition into DST (in Spring) reduces robberies and rape by almost 7% each[1] in their study sample in the US; but has no effect on aggravated assault and murder. On the other hand, transition out of DST during Fall has no statistically significant effect on robberies, rape, aggravated assault and murder.

It is very unlikely that results from Doleac and Sanders (2015) hold for Mexico, especially for homicides, because of the history and severity of violence and the low clearance rate relative to the US.

1.4 Data

Table 1.1 provides average number and likelihood of occurrence of homicides and DTO related homicides per municipality per day. It also breaks these statistics by pre- and post-DST transition during the first six months (labeled as *Spring*) and the last six months (labeled as *Fall*) of a year for the study duration. Sources and brief descriptions of the main variables of interest are given below.

1.4.1 Homicides

Data on homicides are from the Mexican National Health Information System (SINAIS) (Sistema Nacional de Informacion de Salud) which registers daily death certificates. These data are available from 1998-2012 and include information on cause, location (state, municipality and locality), and date of death. The original data was cleaned and death codes standardized with the World Health Organization (WHO) Global Burden of Disease codes by the Center

[1]Point estimates for robberies and rape in Doleac and Sanders (2015) are -0.215 and -0.119 instances per 1000,000, and are significant at the 10% level.

for US-Mexican Studies at University of California, San Diego. I aggregate these data at the municipality-day level (instead of finer locality-day level) for consistency across other data sets and available GIS coordinates.

I am primarily interested in *homicides*, the data for which I get from SINAIS by defining it is deaths that are caused by ”intentional injuries” that are not ”self-inflicted”. From 1998 until 2012 there are 217,585 homicides which account for 3.74% of all deaths in that duration. These occur at an average of 0.0168 occurrences per municipality per day.

1.4.2 DTO Related Homicides

Data on DTO related homicides come from the Center for Research and Teaching in Economics (CIDE) (Centro de Investigacion y Docencia Economicas) Drug Policy Program (PPD) (Programa de Politica de Drogas)[2]. This contains recorded violent events that took place in the context of the war on drugs during the government of Felipe Calderón between December 2006 and November 2011. This has been cleaned to provide data on homicides from three types of violent events at the municipality-day level: (a) executions, defined as intentional homicides that involve an alleged member of a criminal organization (either as a victim or perpetrator) which is not as a result of a prior event and did not involve Mexican authorities; (b) aggressions, defined as an attack by criminal organizations against government institutions or public officers; and (c) confrontations, defined as violent acts which involve clashes with the military, public forces or other criminal organizations (Atuesta et al., 2019). According to this data, homicides from these three events account for almost 87.3%, 1.8% and 10.9% of DTO related homicides respectively. For the purpose of studying the effect of DST transitions on crime, I exclude homicides from (c) and only consider (a) and (b) as DTO related homicides.

From December 2006 until November 2011 there were 42,775 DTO related homicides -

[2]This data was anonymously delivered to CIDE-PPD and is said to have been compiled during President Felipe Calderon’s term to track the war on drugs. Atuesta et al. (2019) provides details on cleaning and validation processes, and potential biases.

which are nearly 50% of all homicides in the municipalities and time period for which CIDE-PPD data are available. These homicides occur at an average of 0.0162 instances per municipality per day.

It must be noted that these numbers represent an underestimate of the total number of homicides and DTO related homicides as these come from registered deaths. There may be a substantial number of undiscovered homicides from unresolved abductions given the history of violence in Mexico. Indeed, officials estimate more than 60,000 missing persons since the 1960s, 97% of which have been reported since the beginning of the war on drugs in 2006[3]. However, without data on abductions, it is difficult to say how this affects my findings.

1.4.3 Sunset Times

Data on exact dates of transitions from Standard Time to DST in Spring and from DST to Standard Time in Fall for the duration of my study period, 1998-2012, were obtained from *www.timeanddate.com*. Sunset times for each municipality-day are calculated using solar mechanics algorithms from Meeus (1991) using municipality centroids.

1.5 Empirical Strategy

I employ a regression discontinuity (RD) design to identify the impact of DST transitions on homicides and DTO related homicides, where the running variable is days before and after DST transitions. These transitions are from Standard Time to DST during the Spring and back from DST to Standard Time in Fall. This differs from using day-of-year as the running variable because DST transitions do not happen on fixed calender days across years. Instead, they take place on specific Sundays in April and October.

As DST always begins and ends on Sundays and crimes are usually higher during week-

[3]See the following Washington Post article: https://wapo.st/2zXBFXS

ends, I control for day-of-week fixed effects. I also include municipality-by-year fixed effects to account for baseline differences in crime across municipalities and years. The exact empirical specification is given by,

$$Y = \alpha + \beta_1 Trans + f_l(DaystoTran) + g_r(DaystoTran) + \gamma_{dow} + \lambda_{mun \times year} \tag{1.1}$$

Where Y are outcomes of interest *homicides* and *DTO related homicide*. $Trans$ is an indicator for DST transitions, i.e. from Standard Time to DST in Spring and from DST to Standard Time in Fall. $DaystoTran$ is days before and after DST transition - the running variable. f_l and g_r are quadratic functions of the running variable to the left and right of the threshold respectively. γ_{dow} and $\lambda_{mun \times year}$ are day-of-week and municipality-by-year fixed effects respectively. I do not control for population as it should be captured by municipality-by-year fixed effects.

I opt to use a 5-week period before and after the transition as bandwidth. I show in the Appendix that results are robust to different bandwidths, including optimal bandwidths given by Calonico et al. (2014).

Finally, standard errors are clustered at municipality and week-of-year to account for variation in crime within municipalities across time, and seasonal variation in crime across municipalities. I show in the Appendix that significance of estimates does not change when clustering standard errors at a more conservative municipality and month-of-year.

1.6 Results

Results for the effect of DST transitions on homicides and DTO related homicides are presented separately for Spring and Fall. These are followed by a discussion on mechanisms and robustness checks. To illustrate RDs, I residualize outcomes by day-of-week and municipality-by-year fixed effects, regress them on a second-order polynomial of the running variable over a 35-day bandwidth, and then plot the predicted values and corresponding confidence intervals along with the residualized outcome averaged over seven day intervals.

1.6.1 Spring Transition

The transition from Standard Time to DST in Spring leads to an increase in the time when the sun sets by one hour on the clock. This is presented in figure 1.4 where I plot the average sunset time leading up to and following the Spring DST transition. This shows that daylight saving increases the average sunset time from almost 6:45 pm to 7:45 pm. Figure 1.5 illustrates the overall RD for homicides and DTO related homicides and shows no jump at the DST transition date. The regression estimate for both of these outcomes (columns (2) and (5) of table 1.2 respectively) show a statistically insignificant effect size of 0.00014 and -0.0008 which account for 0.65% and -5% of pre-DST mean.

Given the concentration of violence crimes in urban areas (Dell, 2015), I conduct the RD analysis separately for rural and urban-only municipalities which is presented in figure 1.6. Since municipalities may have both rural and urban localities, I define municipalities as rural if at least 25% of the municipality population resides in rural areas, and urban-only if 0% of the municipality population resides in rural areas. Panel 1.6b shows a slight positive jump at the DST transition date for homicides in urban-only municipalities; however, this is not statistically significant as shown by corresponding regression estimates in column (4) of table 1.2. Estimates for DTO related homicides are very noisy relative to those for homicides. The point estimate in rural municipalities is 0.0039 which accounts for almost 138% of the pre-DST mean but is statistically insignificant at the 10% level (column (6) of table 1.2). This can be attributed to CIDE-PPD data being available for a shorter duration, and to DTO related homicides being much more volatile.

1.6.2 Fall Transition

The transition from DST to Standard Time in Fall leads to a decrease in the average sunset time from 7 pm to 6 pm (figure 1.7). Figure 1.8 shows a positive jump in homicides after the transition. This corresponds to a statistically significant effect size of 0.0017 which is a 7.76% increase in the DST mean (note that DST is the pre-transition state in Fall) (column (2) of table 1.4). Even though this treatment effect is local to the day of the transition, it equates to nearly 4 additional homicides per year on average. This potentially represents a lower-bound of the overall treatment effect as the effect of the Fall DST transition are likely to last longer than a day.

The positive treatment effect persists when seen separately for rural and urban-only municipalities, the point estimate for which are 0.00081 and 0.0025 respectively (columns (3) and (4) of table 1.4). The higher estimate for urban-only municipalities accounts for a smaller share of DST mean (approximately 16%) relative to that of rural municipalities (54%) but is more precise with statistical significance at the 1%-level relative to significance at the 10%-level for rural municipalities.

As with DST transition in Spring, we do not see effects for DTO related homicides. The RD estimate is for the Fall transition is 0.00062 which is almost 4% of the DST mean and statistically insignificant (column (5) of table 1.4).

1.6.3 Mechanisms

As stated before, I'm unable to disentangle the effect of changing ambient light on factors affecting the expected cost of crime. However, I can explore how effects of changing ambient light vary with a crucial contributor to potential detection and supply of victims - commuting. I do this by studying treatment heterogeneity by day of the week (weekday vs. weekend) and pre-transition sunset time (peak vs. off-peak commuting times) for Spring and Fall transitions. These are estimated using equation 1.1 for the relevant subsample.

Table 1.3 gives results for the effect of Spring DST transitions on homicides and DTO related homicides on weekdays and weekends. As before, estimates for DTO related homicides remain statistically insignificant. However, unlike before, Spring DST transitions reduce homicides on weekends by nearly 8%. Assuming no work related commuting and hence low supply of potential victims on weekends, this result seems to be driven by increase in the probability of detection and capture. Corresponding results for Fall (table

Table 1.6 gives RD estimates for the effect of DST transitions on homicides in urban municipalities by day of the week - weekday vs. weekend - and pre-transition sunset time - peak vs. off-peak commuting hours. I limit this analysis to urban municipalities as people here are more likely to adhere to conventional formal work and, hence, commuting hours. Pre-transition (pre-DST for Spring transitions and DST for Fall transitions) sunset times are considered peak commuting hours if they fall between 5pm and 7pm, and off-peak if they fall between 7pm and 8pm during Spring and 7pm and 10pm during Fall. Due to the location of the sun relative to earth during Spring, there are no geographical locations in the study sample where the pre-transition sunset time in Spring is after 8pm. Estimates for the Spring transition (given in panel 1.6a) are statistically insignificant. Estimates for the Fall transition (given in panel 1.6b) are positive and statistically significant at the 5% level only for peak commuting hours on weekdays (almost 30%) and weekends (almost 39%). This shows that Fall transitions have a disproportionately larger effect in areas where sun sets much earlier implyiing lower probability of detection and higher supply of potentional victims.

1.6.4 Robustness Checks

I discuss and provide a variety of results for robustness checks and placebo tests in the Appendix, and provide a short summary here.

I check the robustness of my results to: (i) changes in the base specification, (ii) changes in the analysis sample, and (iii) changes in bandwidth. I do not conduct every robustness check for DTO related homicides as I find no effects for this outcome in my main results. For changes in base specification (i.e. (i)), I add week-of-year fixed effects, controls for national holidays and observances, and cluster standard errors at a conservative municipality-by-month level in separate regressions. To check if results for homicides hold for subsamples of the analysis sample (i.e. (ii)), I conduct the analysis while limiting the SINAIS data to the years, and then municipalities and years for which I have CIDE-PPD data (i.e. data on DTO related homicides). Finally, for (iii), I check whether results for homicides are robust to a bandwidth choice of 21, 28 and 42 days around the threshold. My results are largely robust to all this checks and point estimates mostly vary within a standard deviation of my main results.

I conduct two falsification exercises or placebo tests. First, I use data from municipalities in states and time periods where DST is not practiced and conduct the RD analysis by assigning DST start and end dates that correspond to the rest of the country in every year. Unfortunately this test is not very informative because coefficients from this test are (mostly) larger than their counterpart in the main analysis but are not statistically significant. The high standard errors of these estimates can be attributed to the small cross-sectional sample that is being used for the test. My second test is much more informative. I only do this test for homicides in Fall transition as this is my main significant result. For this test I conduct RD analysis for my main analysis sample assuming “false” DST transition dates for every day four weeks prior and post the true DST transition date. If effects on homicides are truly driven by the Fall DST transitions, then we should not expect to see statistically significant effects in the same direction on days slightly further away from the true DST transition date; which is indeed what this test shows.

1.7 Conclusion

I show that DST transitions, especially those in Fall, have a statistically and economically significant impact on homicides in Mexico. Earlier sunset times from transition out of DST during

Fall increase homicides by almost 7.8%. I am unable to parse exact mechanisms, but do show evidence suggesting the involvement of these mechanisms using variation in pre-transition sunset times relative commuting hours.

IEP (2013) estimates the social cost of a homicide in Mexico to be 34,776,464 *pesos* (US \$2.6 million) in 2013 Mexican *pesos*[4]. Back-of-the-envelope calculations show that at the very least Fall transitions have a social cost of 145.2 million Mexican *pesos* per year (US \$10.8 million) in 2013 Mexican *pesos*[5], 70% of which is borne by urban municipalities. Since RD estimates give the immediate effect of Fall transitions on homicides, this cost is a lower bound and may significantly underestimate true costs depending on the severity of delayed effects of Fall DST transitions.

It is unclear whether perpetrators are rescheduling their criminal activity from days just before DST transitions or whether these are "new" homicides that would not have occurred in the absence of earlier sunset times. If one categorizes DTO related homicides as organized and premeditated crime, especially during the war on drugs, and the remaining homicides as unpremeditated or not part of organized crime, then my results allude to the latter. Null results (albeit imprecise) for DTO related homicides suggest that changes in ambient light relative to the clock have negligible effects on the factors associated with the expected cost of committing crime for members of organized crime. Combined with higher overall homicides during earlier sunset times, my results suggests that ambient light affects factors associated with the expected cost of committing crime for unorganized criminals who might not have timed the crime in accordance with the DST transition.

[4]IEP (2013) calculates this using estimates of the indirect social cost of murder from McCollister et al. (2010). The estimate from McCollister et al. (2010) - US \$8,442,000 in 2008 dollars - is calculated for the US and equals the mean value of statistical life. To make this relevant for Mexico, IEP (2013) scales this by the ratio of Mexico's GDP per capita relative to the US before converting to 2013 Mexican *pesos*.

[5]These calculations are based on estimated increase in homicides due to Fall DST transition (from table 1.4) and number of municipalities in the overall sample: $0.0017 \times 2{,}456 = 4.18$ homicides per year.

Figures and Tables

Figure 1.1: Timezones in Mexico

Figure 1.2: Intentional homicides in Mexico from 1998 - 2012

Notes: Data downloaded from World Bank Open Data and originally sourced from the UN Office on Drugs and Crime's (UNODC) International Homicide Statistics database.

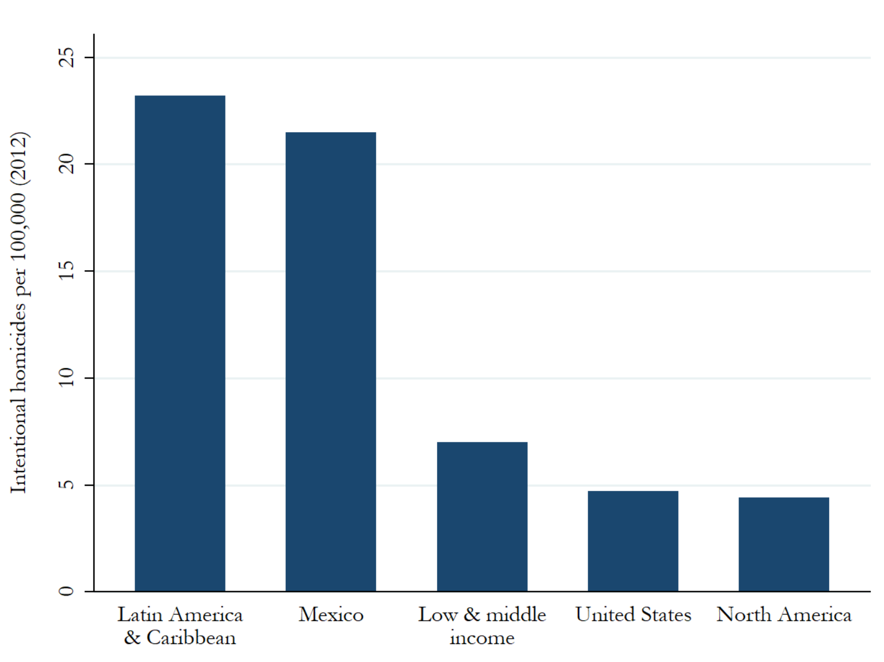


Figure 1.3: Intentional homicides in 2012 for select countries and regions

Notes: Data downloaded from World Bank Open Data and originally sourced from the UN Office on Drugs and Crime's (UNODC) International Homicide Statistics database.

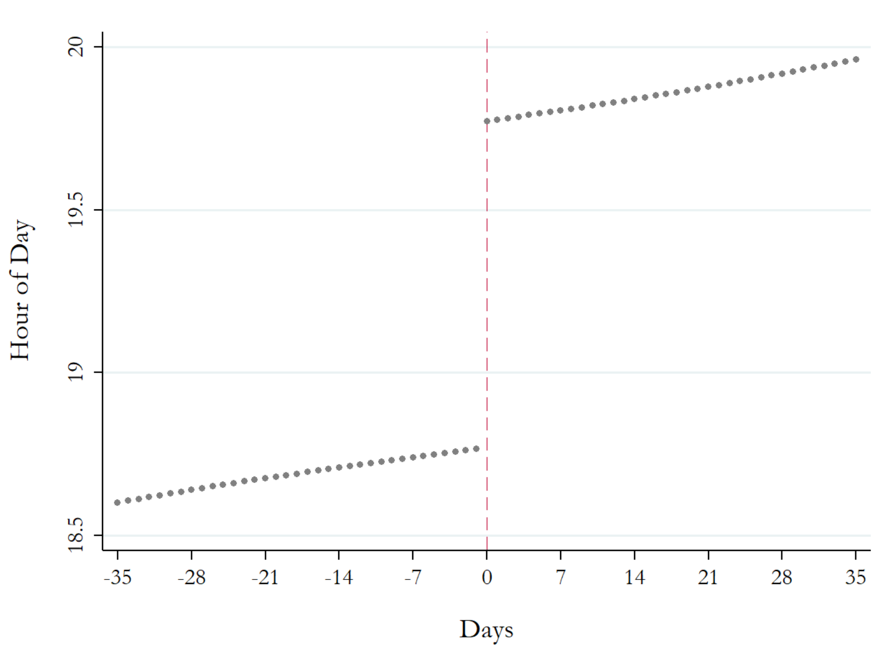
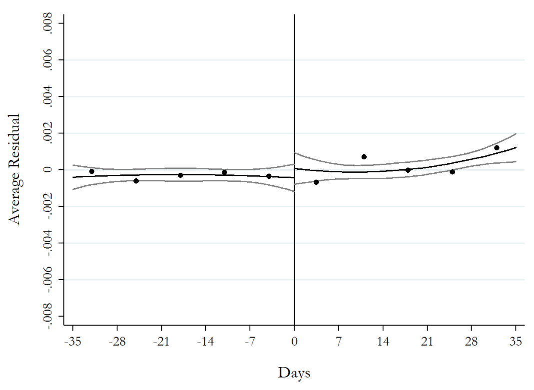
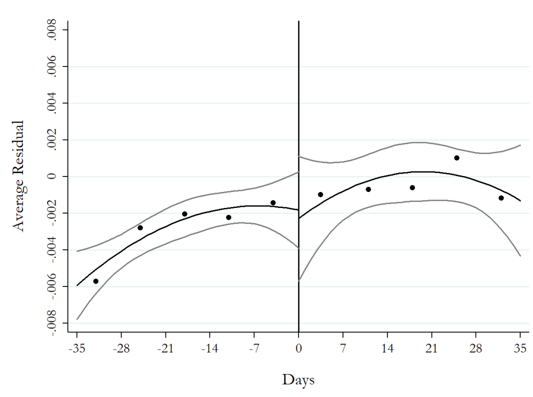


Figure 1.4: Change in sunset time during spring DST transition

Notes: Sunset times for each municipality centroid and day are calculated using the solar mechanics algorithms from Meeus (1991). Each point represents the average sunset time for a given day before and after DST starts in Spring (first week of April) from 1998 - 2012.



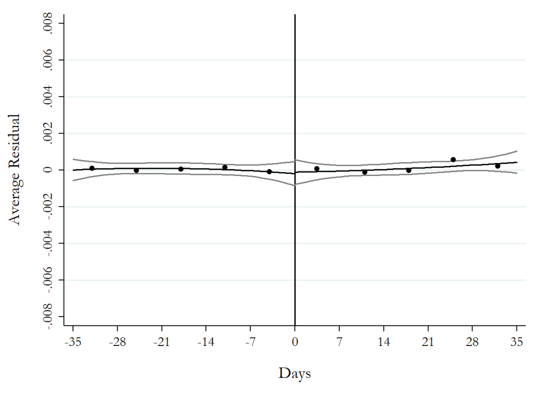
(a) Homicides



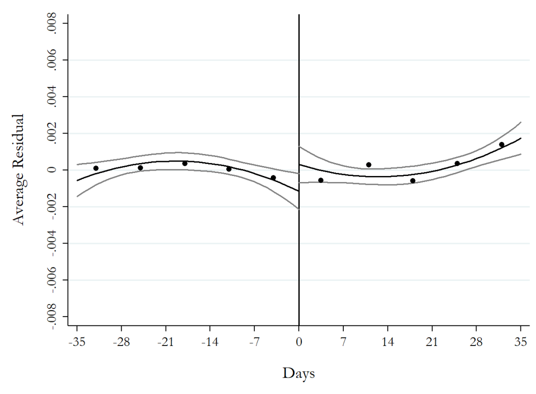
(b) DTO homicides

Figure 1.5: Effect of spring DST transition on homicides and DTO related homicides

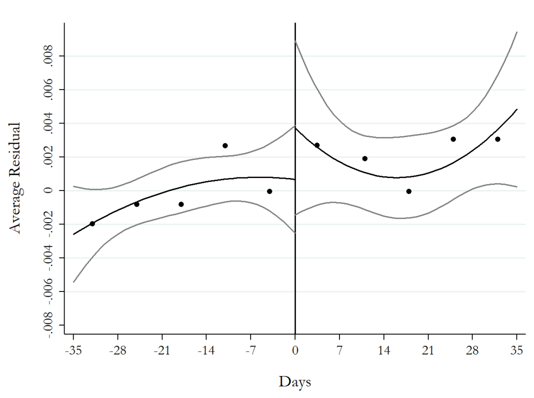
Notes: Outcomes are number of crimes at the municipality-day level. Data for homicides comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for DTO related homicides comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. "Days" refers to days since DST transition takes place. Residuals are generated from a regression of the outcome variable on day-of-week and municipality-by-year fixed effects. Each point in a given plot is the residual, calculated within the corresponding analysis sample, averaged over seven day intervals before and after the DST transition date. Dark solid lines are predicted outcome values based on second-order polynomial regressions, and light solid lines are confidence intervals of the prediction.



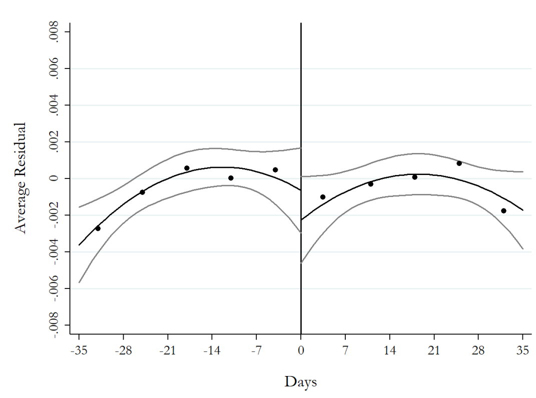
(a) Homicides - rural municipalities



(b) Homicides - urban municipalities



(c) DTO homicides - rural municipalities



(d) DTO homicides - urban municipalities

Figure 1.6: Effect of spring DST transition on homicides and DTO homicides by municipality type

Notes: Outcomes are number of crimes at the municipality-day level. Data for homicides comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is “intentional injuries not caused by self”. Data for DTO related homicides comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. “Days” refers to days since DST transition takes place. Municipalities are categorized as rural if 25% or more of the municipality’s population is rural, and urban if 0% of the municipality’s population is rural. Residuals are generated from a regression of the outcome variable on day-of-week and municipality-by-year fixed effects. Each point in a given plot is the residual, calculated within the corresponding analysis sample, averaged over seven day intervals before and after the DST transition date. Dark solid lines are predicted outcome values based on second-order polynomial regressions, and light solid lines are confidence intervals of the prediction.

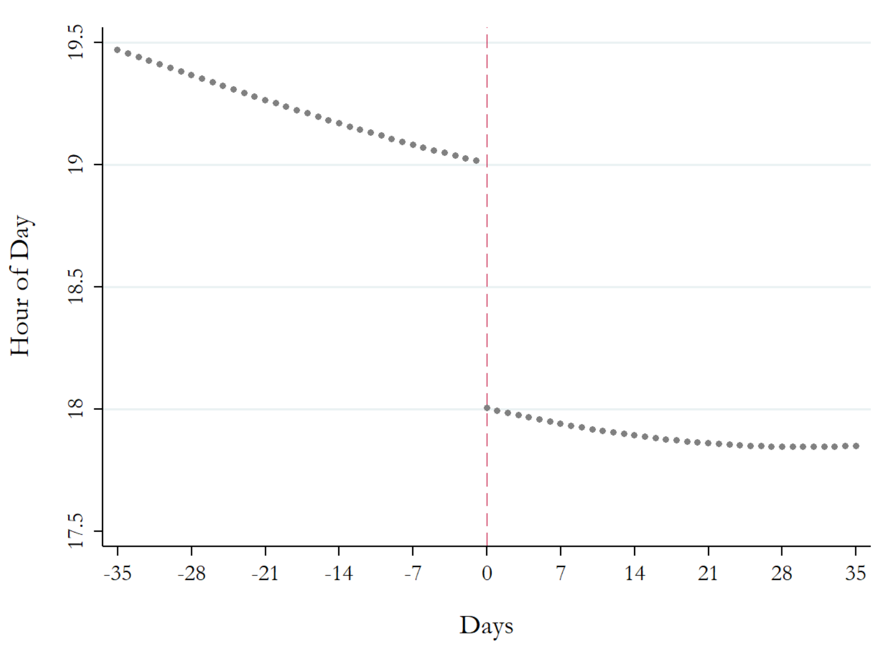
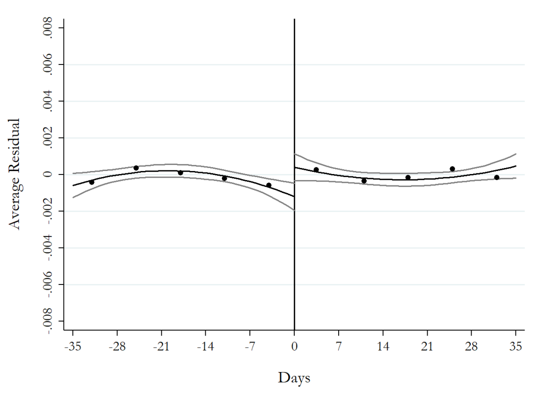
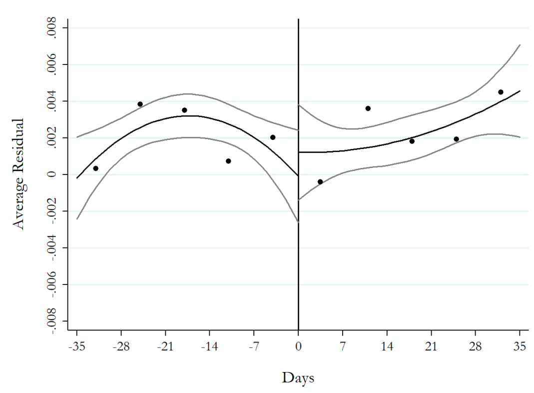


Figure 1.7: Change in sunset time during fall DST transition

Notes: Sunset times for each municipality centroid and day are calculated using the solar mechanics algorithms from Meeus (1991). Each point represents the average sunset time for a given day before and after DST ends in fall (last week of October) from 1998 - 2012.



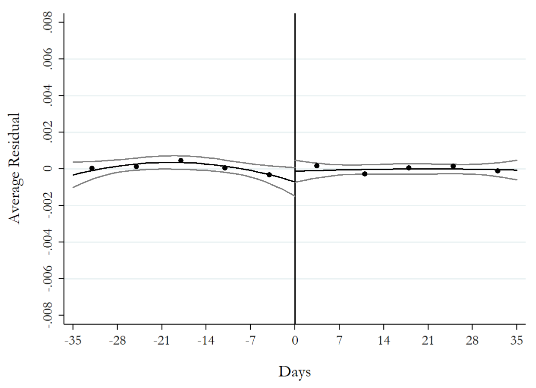
(a) Homicides



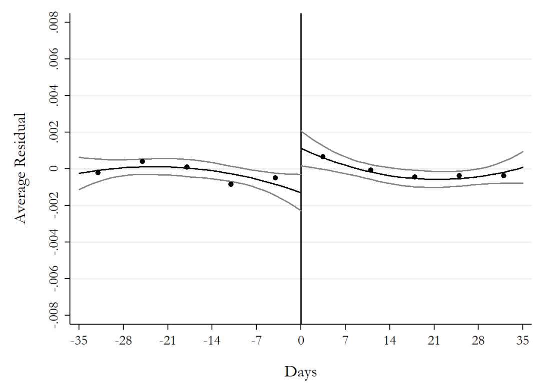
(b) DTO homicides

Figure 1.8: Effect of fall DST transition on homicides and DTO homicides

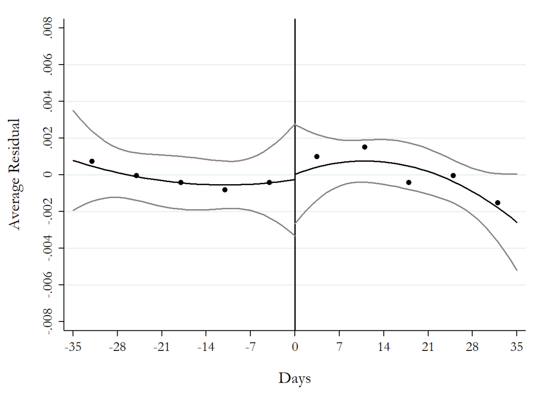
Notes: Outcomes are number of crimes at the municipality-day level. Data for homicides comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for DTO related homicides comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. "Days" refers to days since DST transition takes place. Residuals are generated from a regression of the outcome variable on day-of-week and municipality-by-year fixed effects. Each point in a given plot is the residual, calculated within the corresponding analysis sample, averaged over seven day intervals before and after the DST transition date. Dark solid lines are predicted outcome values based on second-order polynomial regressions, and light solid lines are confidence intervals of the prediction.



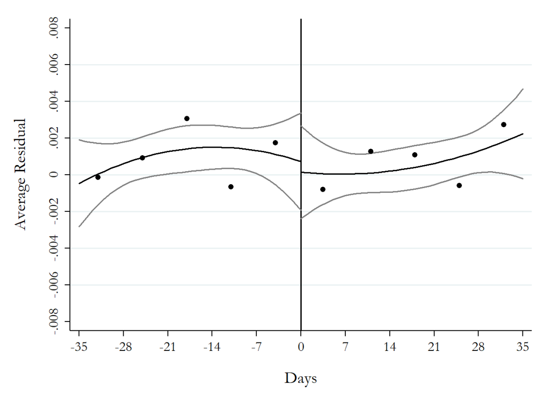
(a) Homicides - rural municipalities



(b) Homicides - urban municipalities



(c) DTO homicides - rural municipalities



(d) DTO homicides - urban municipalities

Figure 1.9: Effect of fall DST transition on homicides and DTO homicides by municipality type

Notes: Outcomes are number of crimes at the municipality-day level. Data for homicides comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for DTO related homicides comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. "Days" refers to days since DST transition takes place. Municipalities are categorized as rural if 25% or more of the municipality's population is rural, and urban if 0% of the municipality's population is rural. Residuals are generated from a regression of the outcome variable on day-of-week and municipality-by-year fixed effects. Each point in a given plot is the residual, calculated within the corresponding analysis sample, averaged over seven day intervals before and after the DST transition date. Dark solid lines are predicted outcome values based on second-order polynomial regressions, and light solid lines are confidence intervals of the prediction.

Table 1.1: Summary table

(a) Homicides

	Total	Spring		Fall	
		Pre-transition	Post-transition	Pre-transition	Post-transition
Number of crimes	.0168	.016	.0172	.0169	.0172
	(.00005)	(.00009)	(.00011)	(.00009)	(.00012)
Probability of crimes	.013	.0128	.0133	.0129	.0132
	(.00003)	(.00006)	(.00007)	(.00006)	(.00007)
Proportion urban	.787				
Average population	37772				
Years	1998-2012				
Observations	12983761	3383390	3059907	4173522	2366942

(b) DTO homicides

	Total	Spring		Fall	
		Pre-transition	Post-transition	Pre-transition	Post-transition
Number of crimes	.0162	.0132	.0169	.0182	.0162
	(.00017)	(.00025)	(.00038)	(.00033)	(.00040)
Probability of crimes	.00903	.00801	.0092	.00985	.00878
	(.00006)	(.00012)	(.00013)	(.00011)	(.00014)
Proportion urban	.899				
Average population	71114				
Years	2006-2011				
Observations	2349518	599215	568262	766652	415389

Notes: This table gives summary statistics for the number and probability of occurrence of homicides (panel 1.1a) and DTO related homicides (panel 1.1b). *Spring* and *Fall* are defined as the first and last 6 months of a year respectively. Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for *DTO related homicides* comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. *Proportion urban* and *Average population* are the proportion of municipalities that are urban and average municipality population respectively calculated over the sample and period the corresponding crime variables are available for.

Table 1.2: RD estimate of the effect of spring DST transition

	Sunset time	Homicides			DTO homicides		
	Overall	Overall	Rural	Urban	Overall	Rural	Urban
DST transition	1***	.00014	.00022	.0012	-.0008	.0039	-.0012
	(.000)	(.001)	(.000)	(.001)	(.002)	(.003)	(.002)
Percentage of control	5.34	.65	10.78	7.37	-5.00	138.19	-10.36
Adjusted R^2	1.00	.18	.01	.14	.13	.01	.12
N	2523766	2523766	297774	1101707	457453	26128	231602

Notes: Each column represents estimates from a single regression estimated from equation 1.1. Data are at the municipality-day level. *Sunset times* are calculated using the solar mechanics algorithms from Meeus (1991). Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for *DTO related homicides* comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. Municipalities are categorized as *rural* if 25% or more of the municipality's population is rural, and *urban* if 0% of the municipality's population is rural. *DST transition* is an indicator for transition into DST during spring (first week of April) which results in the clock shifting forward by one hour. Standard errors are clustered at the municipality X week-of-year level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.3: RD estimate of the effect of spring DST transition by day of week

	Homicides		DTO homicides	
	Weekday	Weekend	Weekday	Weekend
DST transition	.00086	-.0018*	-.0001	-.0017
	(.001)	(.001)	(.003)	(.004)
Percentage of control	5.11	-8.25	-.69	-11.01
Adjusted R^2	.17	.20	.12	.15
N	1777300	746466	322150	135303

Notes: Each column represents estimates from a single regression estimated from equation 1.1. Data are at the municipality-day level. Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for *DTO related homicides* comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. Saturdays and Sundays are categorized into *weekends* and the remaining days of the week make up *weekdays*. *DST transition* is an indicator for transition into DST during spring (first week of April) which results in the clock shifting forward by one hour. Standard errors are clustered at the municipality X week-of-year level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.4: RD estimate of the effect of fall DST transition

	Sunset time	Homicides			DTO homicides		
	Overall	Overall	Rural	Urban	Overall	Rural	Urban
DST transition	-1***	.0017***	.00081*	.0026***	.00062	.0022	-.00032
	(.000)	(.001)	(.000)	(.001)	(.002)	(.003)	(.002)
Percentage of control	-5.24	7.76	54.01	15.96	3.96	94.25	-2.78
Adjusted R^2	1.00	.27	.01	.13	.35	.00	.11
N	2523766	2523766	297774	1101707	452365	25803	228994

Notes: Each column represents estimates from a single regression estimated from equation 1.1. Data are at the municipality-day level. *Sunset times* are calculated using the solar mechanics algorithms from Meeus (1991). Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for *DTO related homicides* comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. Municipalities are categorized as *rural* if 25% or more of the municipality's population is rural, and *urban* if 0% of the municipality's population is rural. *DST transition* is an indicator for transition out of DST during fall (last week of October) which results in the clock shifting back by one hour. Standard errors are clustered at the municipality X week-of-year level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.5: RD estimate of the effect of fall DST transition by day of week

	Homicides		DTO homicides	
	Weekday	Weekend	Weekday	Weekend
DST transition	.0014**	.0019*	-.00086	.0031
	(.001)	(.001)	(.003)	(.003)
Percentage of control	8.77	9.30	-4.68	26.74
Adjusted R^2	.27	.27	.36	.35
N	1777300	746466	317038	135327

Notes: Each column represents estimates from a single regression estimated from equation 1.1. Data are at the municipality-day level. Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for *DTO related homicides* comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. Saturdays and Sundays are categorized into *weekends* and the remaining days of the week make up *weekdays*. *DST transition* is an indicator for transition out of DST during fall (last week of October) which results in the clock shifting back by one hour. Standard errors are clustered at the municipality X week-of-year level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.6: RD estimate of the effect of DST transition on homicides in urban municipalities by day of week and pre-transition sunset time

(a) Spring transition

	Weekdays		Weekends	
	pre-DST Sunset: 5-7pm	pre-DST Sunset: 7-8pm	pre-DST Sunset: 5-7pm	pre-DST Sunset: 7-8pm
DST transition	.0017	.0023	-.00029	-.0015
	(.001)	(.002)	(.002)	(.003)
Percentage of control	12.65	20.38	-1.70	-10.00
Adjusted R^2	.16	.03	.18	.04
N	599750	176100	251895	73962

(b) Fall transition

	Weekdays		Weekends	
	DST Sunset: 5-7pm	DST Sunset: 7-10pm	DST Sunset: 5-7pm	DST Sunset: 7-10pm
DST transition	.0031**	.0013	.0041**	.0028
	(.001)	(.001)	(.002)	(.002)
Percentage of control	29.87	8.84	38.86	14.36
Adjusted R^2	.17	.09	.16	.12
N	351100	424750	147462	178395

Notes: Each column represents estimates from a single regression estimated from equation 1.1. Data are at the municipality-day level. Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Saturdays and Sundays are categorized into *weekends* and the remaining days of the week make up *weekdays*. pre-DST and DST sunset times are the time when the sun sets before DST and during DST, prior to transitions during spring and fall respectively. *DST transition* is an indicator for transition into DST during spring (first week of April) and out of DST during fall (last week of October). Standard errors are clustered at the municipality X week-of-year level. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix: Robustness Checks

In this appendix I provide additional analysis to establish the robustness of results to different specifications, samples and bandwidths. I also conduct falsification tests to check if something other than daylight saving time (DST) transitions are yielding the results that I observe in the main text.

A.1 Specification, Sample and Bandwidths

Table 1.A.1 shows results for different RD specifications for outcome variables homicides and DTO related homicides, and for spring (panel 1.A.1a) and fall (panel 1.A.1b) DST transitions. Columns (1) and (5) replicate original results for reference. Columns (2) and (6) present results from regressions where standard errors are clustered at a more conservative municipality X month-of-year level. Columns (3) and (7) add week-of-year fixed effects to the estimating equation 1.1, and columns (4) and (8) add indicators for national holidays and observances. Results are very stable to the choice of clustering, and relatively stable to the choice of specification. Focusing on homicides during the fall DST transition, controlling for week-of-year and holidays reduces the point estimate; however, these are within one standard deviation of the original estimate and are significant at the 10% level.

Table 1.A.2 presents results for impact on homicides when using different sub-samples for the analysis. Panel 1.A.2a focuses on spring DST transition and panel 1.A.1b focuses on fall DST transition. Columns (1) and (4) replicates original regression results for outcome variables homicides and DTO related homicides respectively for reference. Column (2) is results for homicides using data from years 2006-2011, i.e. the period of time for which I have data available for DTO related homicides. Column (3) presents results when the data is restricted to municipalities and years for which I have data available for DTO related homicides. For both subsets of the data, we see that effect on homicides remain statistically insignificant for

spring transitions (although the sign changes but the absolute value of the estimate is much smaller); whereas, the effect on homicides is higher (although statistically significant at a higher significance level) for fall transitions. This further emphasizes the point that changes in sunset times during fall affect homicides committed by criminals that are not related to drug trade organizations and likely not organized.

Finally, table 1.A.3 presents results with varying bandwidths. For these analyses, I only focus on homicides in the overall sample and homicides in urban-only municipalities. Panel and 1.A.3b show results for spring and fall DST transitions respectively. Columns (1) and (5) replicate original results for the overall sample and urban-only municipality sample respectively (bandwidth of 35 days before and after DST transition dates). Columns (2) and (6), (3) and (7), and (4) and (8) present results for bandwidths of 21 days, 28 days and 35 days before and after DST transition dates respectively. Regardless of bandwidth choice, the effect of spring DST transition on homicides (panel 1.A.3a) remains small and insignificant, although it changes sign in the overall sample. With respect to fall DST transitions (panel 1.A.3a), except for the smallest bandwidth in the urban-only municipality sample, results are within one standard deviation of the original results and are statistically significant at the 10% level.

A.2 Falsification

I conduct two types of falsification tests: in the first test, I estimate the effect of an "assumed" DST transition date on municipalities that do not practice daylight saving; whereas in the second test, I maintain the correct analysis sample, but assign "false" DST transition dates that are at most 28 days before or after the "true" DST transition date.

For the first test, I limit the regression discontinuity (RD) analysis to municipalities in states that do not practice daylight saving[6] and assign them a spring and fall DST transition date

[6]Within the time period for which I have data on homicides (1998-2012), the state of Sonora did not practice daylight saving from 1999-2012, and the state of Quintana Roo did not practice daylight saving in 2012.

for a given year that is prevalent in the rest of the country during that year. Results for this test are presented in table 1.A.4 for spring DST transitions (panel 1.A.4a) and fall DST transitions (panel 1.A.4b). Columns (1) and (3) present results for an unrestricted sample of non-daylight saving practicing municipalities for outcome variables homicides and DTO related homicides respectively; whereas columns (2) and (4) present results for urban-only non-daylights saving practicing municipalities. All point estimates presented in this table are much higher relative to those found in the main results; however, they are all statistically insignificant.

I conduct the second test only for homicides and the fall DST transition as this is the main statistically positive result of the paper. For this test, I maintain the original analysis sample and conduct the RD regressions using "false" DST transition dates that are at most 28 days before or after the "true" DST transition date for a given year. This yields an estimate for each "assumed" DST transition date which I plot in figure 1.A.1 and differentiate statistically significant positive point estimates by marking them with a cross. This graph shows that only two of fifty six point estimates are positive and statistically significant - the day before, and the day of the fall DST transition date; and even between those two, the true fall DST transition date has the larger effect size.

Table 1.A.1: RD estimates with different controls

(a) Spring transition

	Homicides				DTO homicides			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Additional Controls		SEs: MunXmonth	Week-of-year	Holidays		SEs: MunXmonth	Week-of-year	Holidays
DST transition	.00014	.00014	.00021	.0001	-.0008	-.0008	-.00049	-.00096
	(.001)	(.001)	(.001)	(.001)	(.002)	(.003)	(.003)	(.002)
Adjusted R^2	.18	.18	.18	.18	.13	.13	.13	.13
N	2523766	2523766	2523766	2523766	457453	457453	457453	457453

(b) Fall transition

	Homicides				DTO homicides			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Additional Controls		SEs: MunXmonth	Week-of-year	Holidays		SEs: MunXmonth	Week-of-year	Holidays
DST transition	.0017***	.0017***	.0013*	.0011*	.00062	.00062	-.00089	.0016
	(.001)	(.001)	(.001)	(.001)	(.002)	(.002)	(.002)	(.002)
Adjusted R^2	.27	.27	.27	.27	.35	.35	.35	.35
N	2523766	2523766	2523766	2523766	452365	452365	452365	452365

Notes: Each column represents estimates from a single regression estimated from slight variations of equation 1.1. Data are at the municipality-day level. *Sunset times* are calculated using the solar mechanics algorithms from Meeus (1991). Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for *DTO related homicides* comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. *DST transition* is an indicator for transition into DST during spring (first week of April - panel 1.A.1a), and transition out of DST during fall (last week of October - panel 1.A.1b). Columns (1) and (5) for both panels replicate original regressions for spring and fall transitions respectively for reference. Columns (2) and (6) cluster standard errors at municipality X month-of-year level. Columns (3) and (7) add week-of-year fixed effects, and columns(4) and (8) add indicators for national holidays. Standard errors are clustered at the municipality X week-of-year level unless otherwise stated. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.A.2: RD estimates on homicides using different samples

(a) Spring transition

	Homicides			DTO homicides
	(1)	(2)	(3)	(4)
Sample		2006-2011	Overlap	
DST transition	.00014	-.00034	-.00059	-.0008
	(.001)	(.001)	(.003)	(.002)
Adjusted R^2	.18	.19	.19	.13
N	2523766	1013596	457453	457453

(b) Fall transition

	Homicides			DTO homicides
	(1)	(2)	(3)	(4)
Sample		2006-2011	Overlap	
DST transition	.0017***	.0029**	.0043*	.00062
	(.001)	(.001)	(.002)	(.002)
Adjusted R^2	.27	.34	.36	.35
N	2523766	1013596	452365	452365

Notes: Each column represents estimates from a single regression estimated from equation 1.1. Data are at the municipality-day level. *Sunset times* are calculated using the solar mechanics algorithms from Meeus (1991). Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for *DTO related homicides* comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. *DST transition* is an indicator for transition into DST during spring (first week of April - panel 1.A.2a), and transition out of DST during fall (last week of October - panel 1.A.2b). Columns (1) and (4) for both panels replicate original regressions for spring and fall transitions respectively for reference. Column (2) is the regression for *homicides* run on a subset of data from years 2006-2011 (the period of time for which I have data available for *DTO homicides*. Column (3) is the regression for *homicides* run on data restricted to municipalities and years for which I have data available for *DTO homicides*. Standard errors are clustered at the municipality X week-of-year level unless otherwise stated. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.A.3: RD estimates on homicides by different bandwidths

(a) Spring transition

	Overall				Urban			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth	35	21	28	42	35	21	28	42
DST transition	.00014	-.00053	.000092	-.00012	.0012	.00081	.0014	.00038
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Adjusted R^2	.18	.18	.18	.17	.14	.15	.15	.14
N	2523766	1528478	2026122	3021410	1101707	667231	884469	1318945

(b) Fall transition

	Overall				Urban			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth	35	21	28	42	35	21	28	42
DST transition	.0017***	.0018**	.0015**	.0012**	.0026***	.0011	.0019**	.0023***
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Adjusted R^2	.27	.27	.27	.26	.13	.13	.13	.13
N	2523766	1528478	2026122	3021410	1101707	667231	884469	1318945

Notes: Each column represents estimates from a single regression estimated from equation 1.1. Data are at the municipality-day level. *Sunset times* are calculated using the solar mechanics algorithms from Meeus (1991). Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". *DST transition* is an indicator for transition into DST during spring (first week of April - panel 1.A.3a), and transition out of DST during fall (last week of October - panel 1.A.3b). Columns (1) to (4) show results from analysis on an unrestricted sample. Columns (5) to (6) show results from analysis conducted on urban-only municipalities. Standard errors are clustered at the municipality X week-of-year level unless otherwise stated. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.A.4: RD estimates for municipalities that do not practice DST

(a) Spring transition

	Homicides		DTO homicides	
	(1)	(2)	(3)	(4)
Municipality type	Overall	Urban	Overall	Urban
DST transition	.0031	.012	-.00014	.011
	(.004)	(.010)	(.007)	(.013)
Adjusted R^2	.08	.07	.05	.01
N	71308	18961	16330	5751

(b) Fall transition

	Homicides		DTO homicides	
	(1)	(2)	(3)	(4)
Municipality type	Overall	Urban	Overall	Urban
DST transition	.003	-.0015	.009	-.0066
	(.004)	(.009)	(.014)	(.013)
Adjusted R^2	.08	.07	.03	.02
N	71278	18955	16146	5680

Notes: Each column represents estimates from a single regression estimated from equation 1.1 using data from municipalities that do not practice DST. Data are at the municipality-day level. *Sunset times* are calculated using the solar mechanics algorithms from Meeus (1991). Data for *homicides* comes from SINAIS and is a count of registered deaths at the municipality-day level from 1998-2012 where the cause of death is "intentional injuries not caused by self". Data for *DTO related homicides* comes from CIDE-PPD and is a count of homicides related to DTOs, as categorized by key officials of the Mexican government, at the municipality-day level from 2006-2011. *DST transition* is an indicator for the "assumed" transition into DST during spring (first week of April - panel 1.A.4a), and 'assumed" transition out of DST during fall (last week of October - panel 1.A.4b) using transition dates for that year from municipalities that practice DST. Standard errors are clustered at the municipality X week-of-year level unless otherwise stated. Statistical significance is denoted as: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

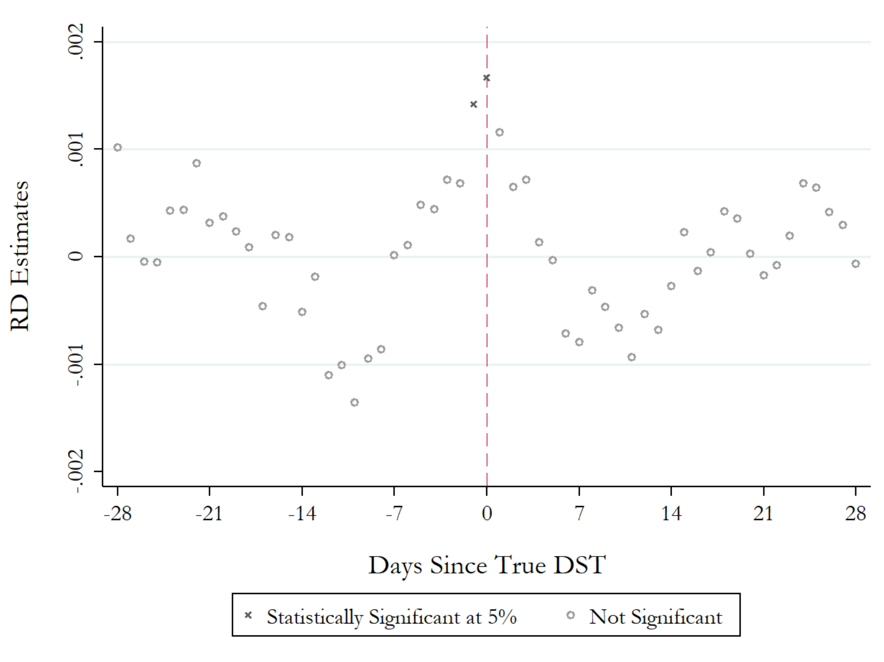


Figure 1.A.1: Scatter plot of estimated DST effect on homicides during fall transitions by days since true DST

Notes: Each point represents estimates from a single regression estimated from equation 1.1 where DST transition is assumed to be a day 4 weeks before or after the "true" fall DST transition date in a given year, and the sample is restricted to 5 weeks before and after the assumed DST transition date (i.e. the bandwidth used in the main analysis). Statistically significant (at the 5% level) positive estimates are marked with a cross, and the remaining estimates are marked with a hollow circle.

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. 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; Vogl, 2013).

The effect of displacement on marriage outcomes is theoretically ambiguous. There are

[1] For example, women who are displaced, whether internationally or internally, are at a higher risk for sexual and gender based violence (UNICEF et al., 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).

reasons to think that displacement could induce earlier marriage for displaced women (Corno et al., Forthcoming; Mourtada et al., 2017; Schlecht et al., 2013; Nour, 2009). For example, 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, marriage market theories suggest that displacement could delay marriage for displaced women (Oppenheimer, 1988). Displacement may increase the search cost of finding good matches if families have less information and limited social networks in their new destinations. Additionally, displaced families may not want their daughters to marry if they are productive household members in home production or the labor market. 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 no 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. 1947-48), 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, 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 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 difference 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 concerns with 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 the main body of this 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 6.2 percentage points more likely to be married at partition (defined as marrying between years 1946 - 1949; comparison mean is 2.3%). 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 of marriage and likelihood of marriage

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

before 18.

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 0.38 more children born over a control mean of 5.64). 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, 3.1 percentage points more likely to have completed primary school, and have more surviving children. The data also allows us to examine spousal quality and 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.

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 (Usta et al., 2008; Masterson et al., 2014; Vu et al., 2014; Keygnaert et al., 2012; Wirtz et al., 2014), the challenges women face in accessing healthcare (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 (Becker et al., forthcoming; Sieverding et al., 2018; Frattini et al., 2020; Maia et al., 2018).[3] Our paper also relates to and is broadly consistent with work examining the links between early marriage and later life outcomes for women (Field and Ambrus, 2008; Delprato et al., 2017; de Groot et al., 2018; Raj et al., 2019; Jayachandran, 2015). 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

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

Mirza, 2019).

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 allow 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.

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

[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

[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.

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 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 their time of survey. Finally, we define non-migrants as those who have never migrated according to their survey response.

Table 2.1a 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 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 to identifying the causal impact of displacement using 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 effect of displacement.

We study recently displaced individuals (those who have been displaced in the 5 years before their time of 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 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 survey. We choose non-migrant

[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. Additionally, we will be using the exogenous nature of partition to obtain causal estimates, while this exercise is purely descriptive.

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 survey. $Displaced_d$ is an indicator of whether the individual was recently displaced (displaced within 5 years of time of 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 X year fixed effect with more granular destination X country X year and birthplace X country X year fixed effects. Results of these robustness checks are reported in Appendix B. Additionally, we report all results with and without clustered standard errors. To report results at the country level, we pool data across years for a given country rather than all datasets and replace country X year fixed effects with year fixed effects in Equation 2.1.

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

[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.

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 2.1a). 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 X country X year (Appendix Figure 2.A.1) and birthplace X country X year fixed effects[9] (Appendix Figure 2.A.2). All cross-country results are also robust to clustering standard errors at the geography X year X displaced level, where geography X year is the level of the fixed effects in the given specification. Appendix Figure 2.A.3 reports the robustness of Figure 2.1 to standard errors clustered at the country X year X displaced level. This pattern is also similar across individual countries. We report the analogous country-level results where we replace country X year fixed effects with year fixed effects in Equation 2.1.

Having shown strong correlations that demonstrate higher marriage rates among young displaced women across different countries and time periods, we now consider the specific case of displacement into Pakistan caused by the partition of India in 1947 to establish a causal interpretation.

2.3 Results from the Partition of India in 1947

2.3.1 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

[9]To include birthplace X country X 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).

to one of the largest and most rapid migrations in human history. It is estimated that about 14.5 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, Hachette UK). 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

[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 UP, 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).

provides a unique opportunity to study the causal impact of displacement without suffering from issues of migration selection and timing in the the cross-country evidence.

Using data from Pakistan's 1973 Housing, Economic, Demographic Characteristics (HED) survey (detailed further in Section 2.3.2) 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.

2.3.2 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 survey) were born in India and have been living in their current locality in Pakistan for 24 to 27 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 we observe 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 2.1a 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 survey (that is, who were born at or before the time of partition) are displaced.

2.3.3 Empirical Strategy

Our identification strategy is a difference-in-difference 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 2.3.2).

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 who were 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 most 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 and 1 to 9 years old at partition.

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. We choose to exclude women over the age of 32 at partition due to concerns of 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

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

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

(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 x urban) fixed effects in Equation 2.2. Our results are very similar under this specification. However, given the spatial distribution of displaced women, it is possible that they may have sorted into location non-randomly. We do not place heavy emphasis on this set of results.

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 2.B.1b, 2.B.2b, and 2.B.3b. 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-difference analysis, but Canay et al. (Forthcoming) 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).

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

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

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 results from this specification for two of our primary marriage outcomes, child marriage and age at first marriage, are reported in Figure 2.6, which shows that the parallel trends assumption is satisfied. Parallel trends tests for other outcomes are reported in Appendix Figure 2.B.1.

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. This increases our confidence that our results are not driven by mortality that is correlated with marriage timing.

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 causal 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, 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 (as in Corno et al. (Forthcoming)) of native born women would lead to a lower bound effect.

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, there is nothing sacred about 1949 as the cutoff point in terms of the robustness of our results. 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 x 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.

2.3.4 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 survey (25-26 years after partition). Overall, 98.9% of women in the main analysis sample have ever been married by the time of 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 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 6.2 percentage points more likely to get married at the time of partition (defined as being married for the first time between years 1946 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).[16]

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 lives in an urban area (Table 2.B.1b). 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). All together, 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 2.4b). 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 2.B.3b).

[16] At the time of survey, women who were aged 1 to 5 years at the time of partition were 26 to 31 at the time of survey. Women in this young group had not all been married by the time of 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).

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. 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%).

The increase in schooling is similar to recent evidence from (Becker et al., forthcoming) who, in the context of the redrawing of Poland's boundaries following World War II, find that displaced people and their descendants invest more in human capital. However, in our context, we find that increased investment in education due to displacement for young females depends crucially on the timing of displacement during their childhood.

Displaced adolescent women also have higher fertility as measured by both total children ever born and children surviving at the time of 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 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.[17]

2.4 Conclusion

Using data from multiple countries over many time periods, our results show that displacement increases the likelihood of early marriage for young women. We do not find the same results for young displaced men. In the specific case of the partition of India, where we can address some of the issues around refugee selection, we further find that displacement and early marriage leads to a lower likelihood of continued education and increases fertility. Early marriage has been linked to lower schooling and higher fertility in a number of papers, and our results are broadly consistent with this.

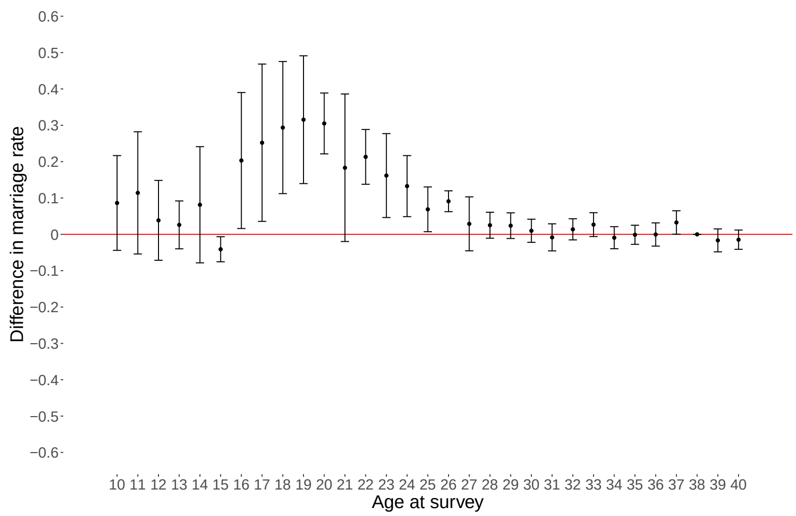
Our main take away is that while displacement events are traumatic in general, they can have particularly negative consequences for women. Certainly, *conditional* on displacement, marriage might have a protective effect if marriage prevents women from suffering violence of other kinds. However, without data on other outcomes faced by women these tradeoffs are impossible to evaluate. From a policy perspective, it seems clear that attention must be paid to women's welfare during periods of displacement, as women might face a Hobson's choice between early marriage and other forms of violence.

Chapter 2 is coauthored with Prashant Bharadwaj and Frances Lu and has been prepared for submission for publication. I am grateful to Frances for her outstanding contribution to this

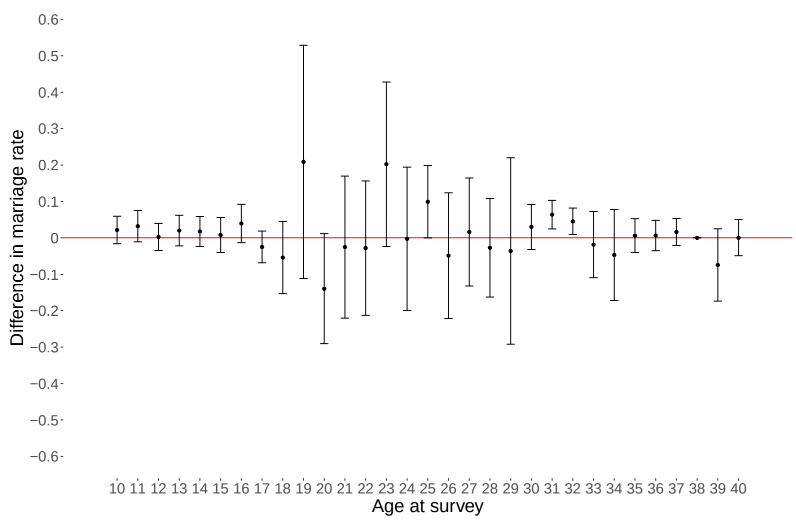
[17]Our estimates may understate the impact of displacement on fertility for two reasons. First, women who were young at partition may not have completed fertility by the time of survey. At the time of survey, women who were ages 1 to 17 at partition are around 26 to 43 at time of survey. This means some women, particularly women in the 1 to 5 age group, may not have completed their fertility. At the very least, our results show young displaced women 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 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-difference estimates for fertility downwards, particularly for the outcome of the number of children born.

chapter, and to Prashant for his guidance and support.

Figures and Tables



(a) Women



(b) Men

Figure 2.1: Cross-country evidence: Effect of displacement on marriage rates

Notes: 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. The analysis sample consists of recently displaced (displaced in last 5 years) and non-migrants of the indicated gender. Data is pooled over all countries and survey years, excluding the 1973 Pakistan dataset (since partition occurred 25-26 years prior to time of survey).

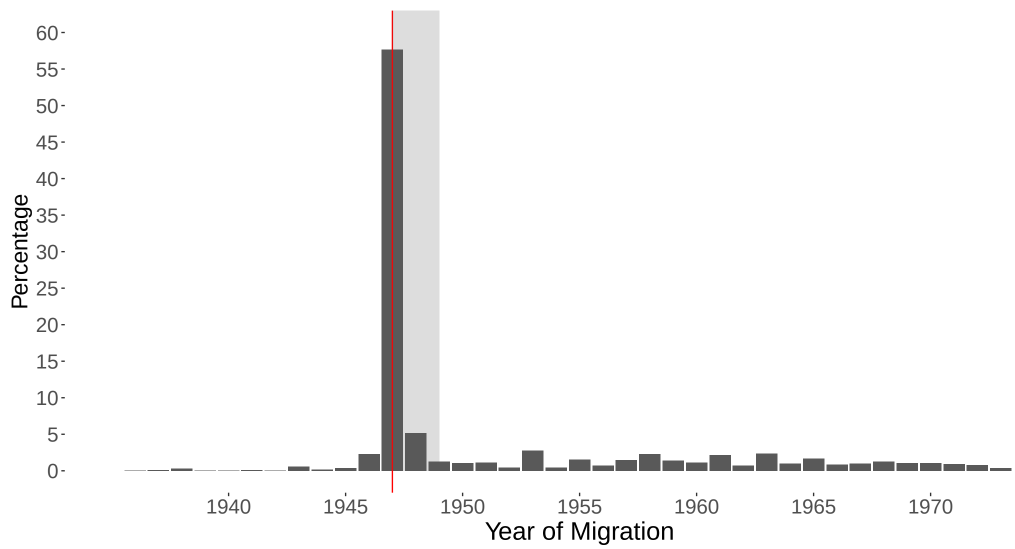
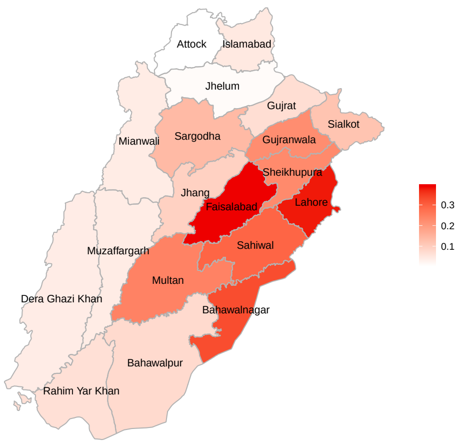
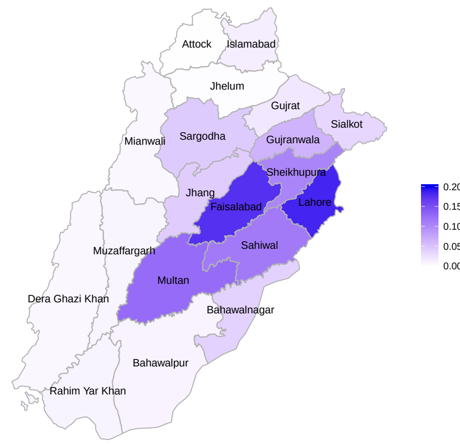


Figure 2.2: Pakistan evidence: Migration timing of Indian-born females

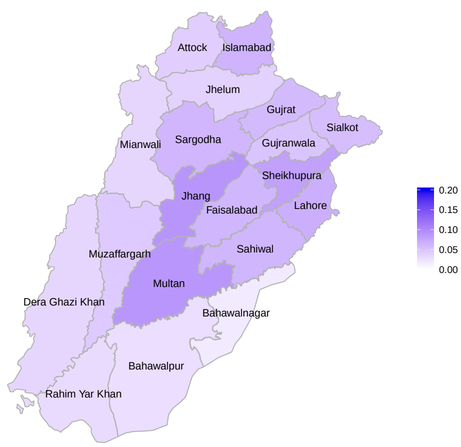
Notes: This figure reports empirical densities of the year of last migration for Indian-born females by age group. The year of migration is calculated as 1973 minus the years since last migration which is reported in the data. Percentages are calculated as 100 times the weighted proportion of individuals from a given age group that migrated in a given year. The proportions are weighted by inverse probability weights. The analysis sample is all Indian-born women alive at the time of partition who are living in Punjab at the time of survey. The data are from the 1973 Pakistan HED survey.



(a) Proportion of women in each district who were displaced



(b) Proportion of displaced women in each district



(c) Proportion of native-born women in each district

Figure 2.3: Pakistan evidence: Spatial distribution of Punjabi female population

Notes: 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. The sample is native-born (born in Pakistan) and displaced (Indian-born females who migrated 24 to 26 years ago) females alive at time of partition who were living in Punjab at the time of enumeration. Panel (b) reports the weighted proportion of displaced women in Punjab who were living in each district. The sample is displaced women who were ages 1 to 32 at time of partition and were living in Punjab at the time of enumeration. Panel (c) reports the weighted proportion of native-born women in Punjab who were living in each district. The sample is native-born women who were ages 1 to 32 at time of partition and were living in Punjab at the time of enumeration. The data are from the 1973 Pakistan HED survey.

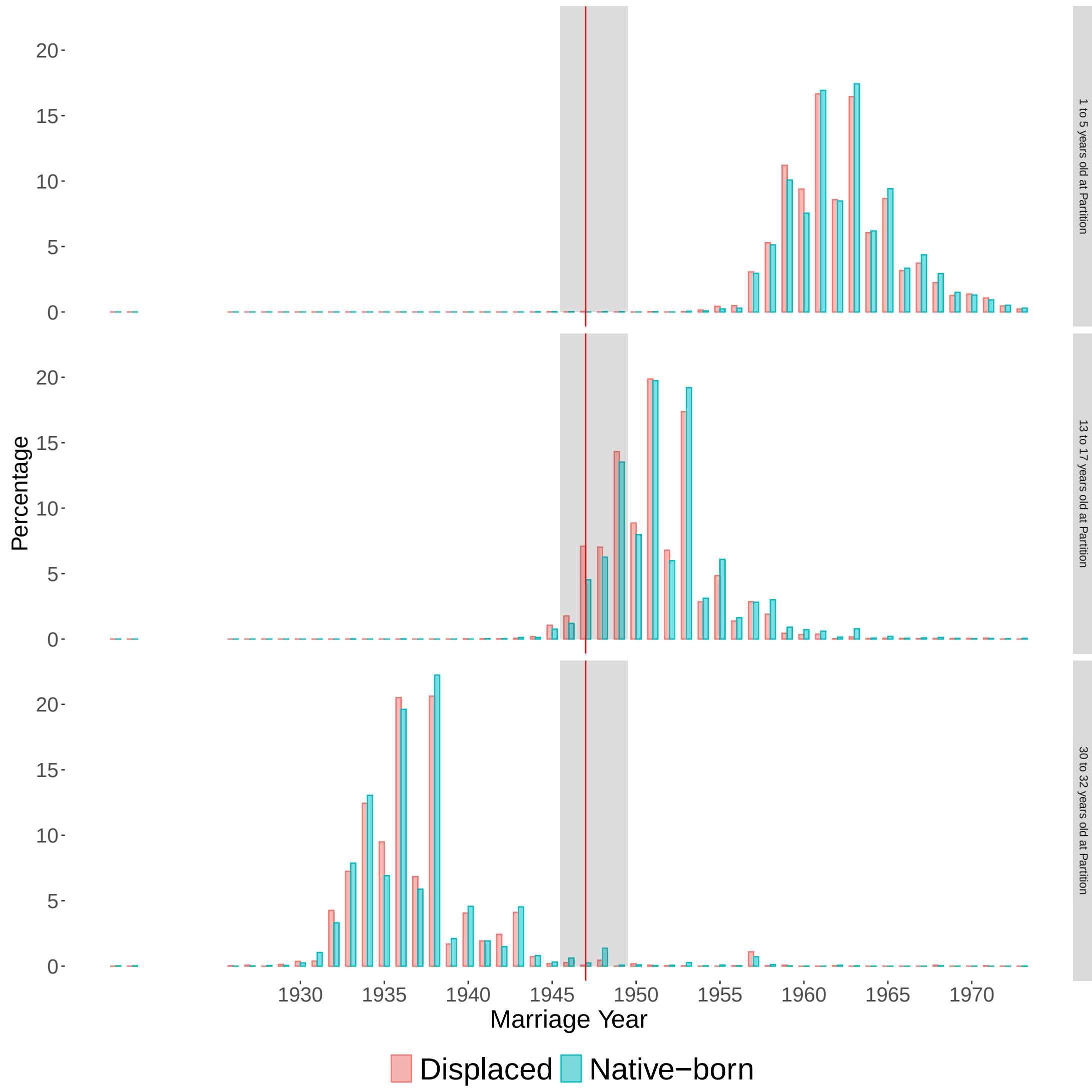


Figure 2.4: Pakistan evidence: Marriage timing by age at partition

Notes: 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. The year of marriage is calculated as 1973 minus the years since first marriage. Years since first marriage is calculated as the difference between current age and age at first marriage. Percentages are calculated as 100 times the weighted proportion of individuals from a given age group that migrated in a given year. The proportions are weighted by inverse probability weights. The analysis sample is native-born (born in Pakistan) and displaced (Indian-born females who migrated 24 to 26 years ago) women who were ages 1 to 5, 13 to 17, or 30 to 32 at the time of partition and were living in Punjab at the time of enumeration. The data are from the 1973 Pakistan HED survey.

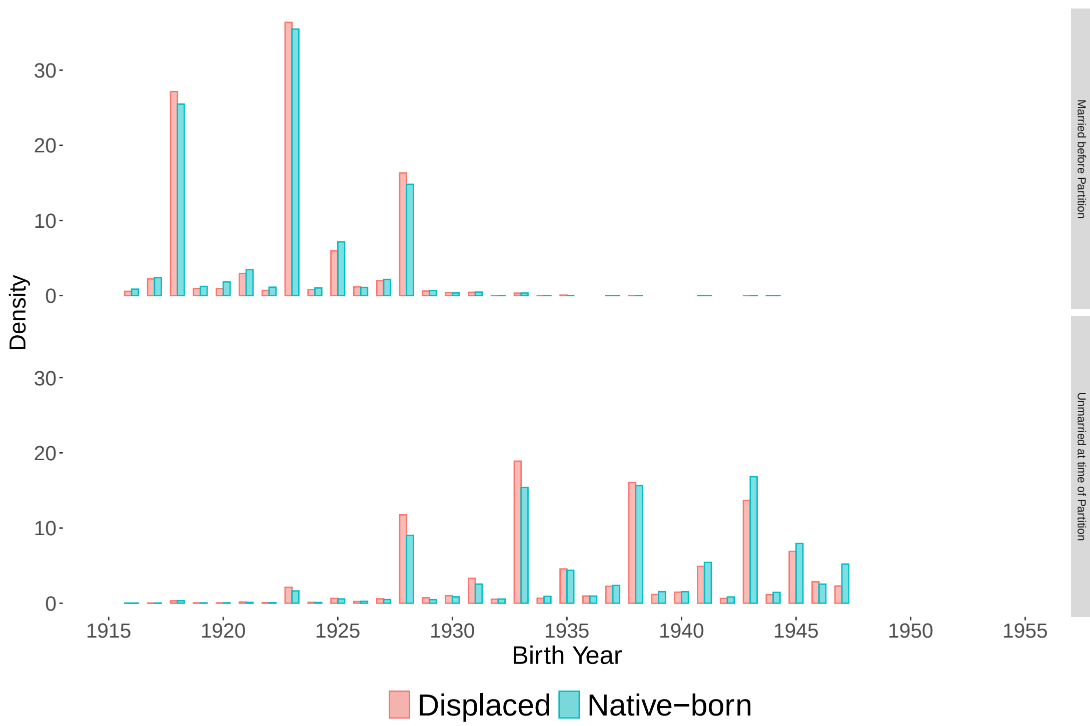
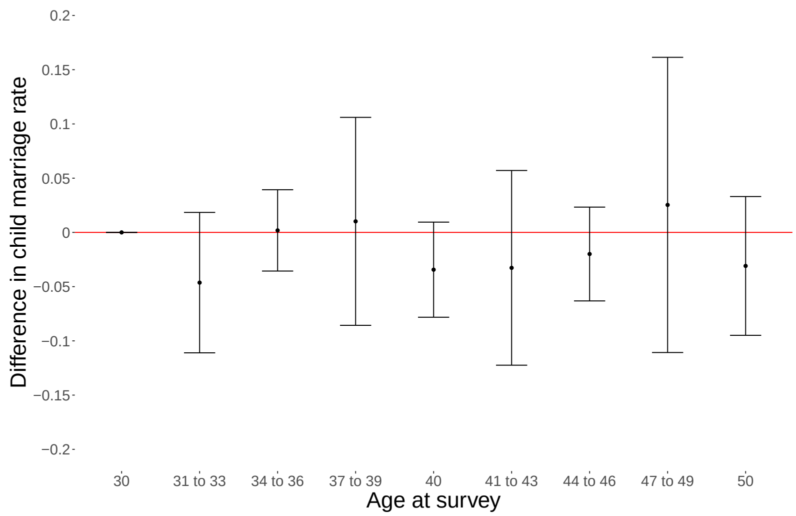
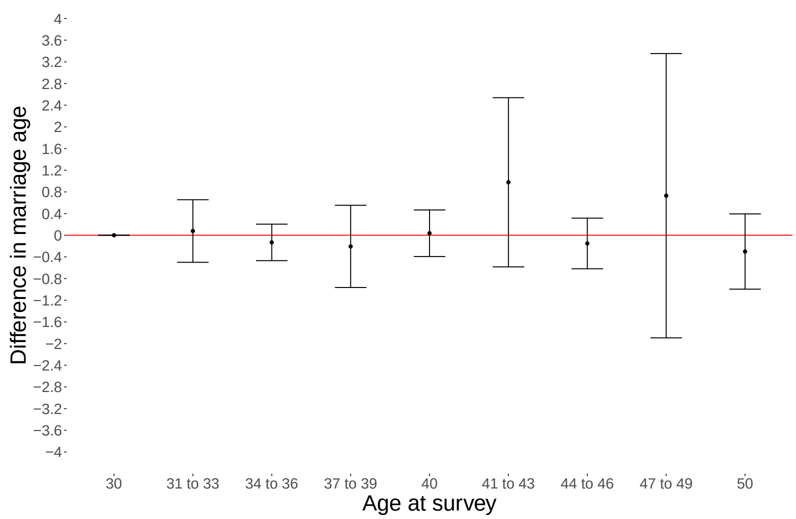


Figure 2.5: Pakistan evidence: Mortality risk at partition by marriage status

Notes: 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. The year of birth is calculated as 1973 minus the respondents reported age. Percentages are calculated as 100 times the weighted proportion of individuals from a given age group that migrated in a given year. The proportions are weighted by inverse probability weights. The analysis sample is native-born (born in Pakistan) and displaced (Indian-born females who migrated 24 to 26 years ago) women who were ages 1 to 32 at the time of partition and were living in Punjab at the time of enumeration. The data are from the 1973 Pakistan HED survey.



(a) Child marriage rate



(b) Marriage age

Figure 2.6: Pakistan evidence: Parallel trends

Notes: This figure plots parallel trends tests for two primary outcomes: child marriage and marriage age. 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. Individual observations are weighted by inverse sampling probability. 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 data are from the 1973 Pakistan HED survey.

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										

Notes: 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 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 recently displaced women. A recently displaced women is defined as a displaced women who was displaced in the last 5 years. The rates for recently displaced women are reported by age group at time of enumeration. Blank boxes indicate that the relevant data was not available for that country. All statistics are weighted using inverse sampling probabilities.

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 at partition
	(1)	(2)	(3)	(4)
Young X Displaced	.01***	.038**	-.28**	.062***
	(.002)	(.016)	(.126)	(.009)
Young	-.013***	.06***	-.63***	.23***
	(.001)	(.008)	(.066)	(.004)
Displaced	.00024	.019	-.25**	-.015***
	(.001)	(.013)	(.112)	(.003)
Comparison mean	.997	.308	19.3	.0231
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 at partition
	(1)	(2)	(3)	(4)
Young X Displaced	-.014***	-.0043	.14	.015***
	(.004)	(.016)	(.124)	(.003)
Young	-.042***	.13***	-1.2***	-.023***
	(.002)	(.008)	(.063)	(.002)
Displaced	.00024	.019	-.25**	-.015***
	(.001)	(.013)	(.112)	(.003)
Comparison mean	.997	.308	19.3	.0231
Adjusted R^2	.01	.01	.02	.02
N	35209	32967	32967	32967

Notes: Each column represents a single regression estimated from Equation 2.2. The data are from the 1973 Pakistan HED survey. 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. "Displaced" is an indicator of whether the respondent was displaced from India around the time of partition (between 1947-1949). 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 at partition" is an indicator variable of whether a respondent was first married between 1946 - 1949. 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 2.3: Pakistan evidence: Age at the time of partition and other 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	2.4e-16*	.34***	-.094
	(.004)	(.003)	(.000)	(.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

Notes: Each column represents a single regression estimated from Equation 2.2. The data are from the 1973 Pakistan HED survey. 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. "Young" is an indicator of whether the respondent was between the ages of 10 and 17 at the time of partition. "Displaced" is an indicator of whether the respondent was displaced from India at the time of partition. 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 "Continued education after partition" is an indicator of whether the respondent's highest level of education attainment occurred after 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 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 2.4: Pakistan evidence: Age at the time of partition and spousal characteristics

(a) Young is 13 to 17

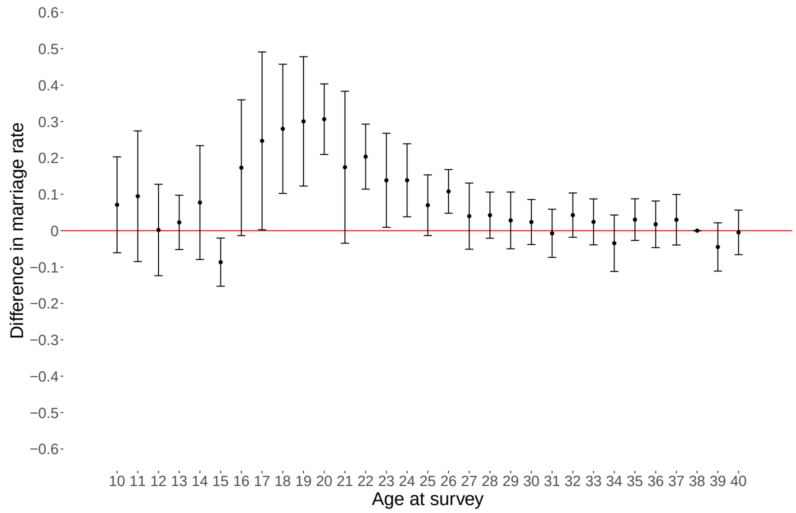
	Age difference	Same nativity	Literate	Completed primary school	Continued education after partition
	(1)	(2)	(3)	(4)	(5)
Young X Displaced	-.13	-.0032	.0026	.0037	.00085
	(.215)	(.006)	(.014)	(.012)	(.002)
Young	.51***	-.011***	.086***	.064***	.013***
	(.107)	(.002)	(.007)	(.006)	(.001)
Displaced	.024	-.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

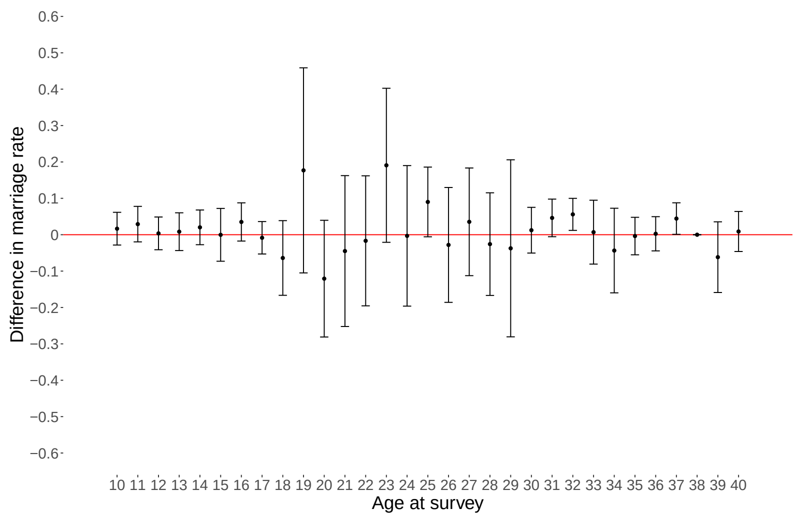
	Age difference	Same nativity	Literate	Completed primary school	Continued education after partition
	(1)	(2)	(3)	(4)	(5)
Young X Displaced	-.58***	.011*	.054***	.057***	.065***
	(.213)	(.006)	(.015)	(.013)	(.008)
Young	-.094	-.048***	.17***	.14***	.17***
	(.101)	(.002)	(.007)	(.006)	(.003)
Displaced	.024	-.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

Notes: Each column represents a single regression estimated from Equation 2.2. The data are from the 1973 Pakistan HED survey. 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. "Young" is an indicator of whether the respondent was between the ages of 10 and 17 at the time of partition. "Displaced" is an indicator of whether the respondent was displaced from India 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. 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: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Appendix A: Robustness Checks for Cross-country Evidence



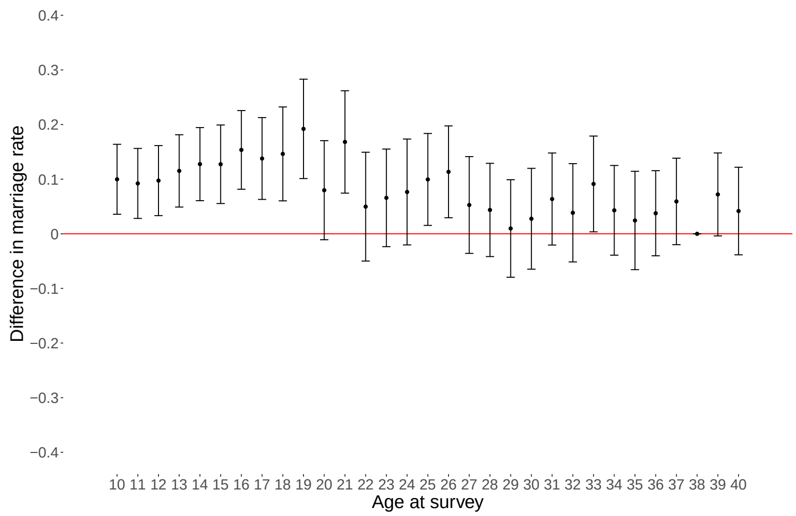
(a) Women



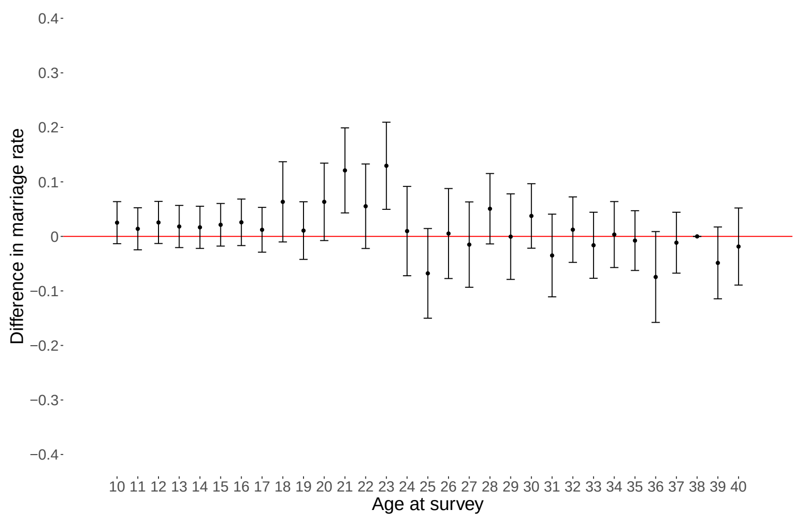
(b) Men

Figure 2.A.1: Cross-country evidence: Effect of displacement on marriage rates (destination fixed effects)

Notes: 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 survey. The omitted age group is age 38 at time of 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. The analysis sample consists of recently displaced (displaced in last 5 years) and non-migrants of the indicated gender and age at survey. Data is pooled over all countries, excluding Pakistan (since partition occurred 26 years prior to survey year).



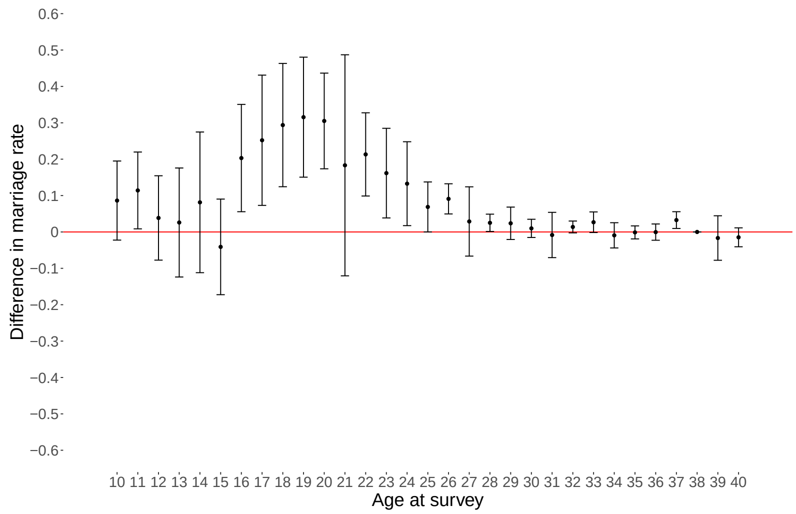
(a) Women



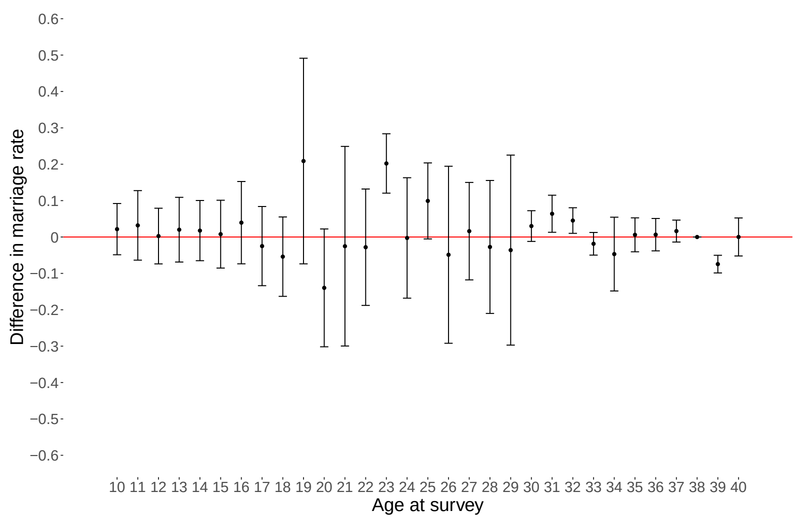
(b) Men

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

Notes: 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 omitted age group is age 38 at time of 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. The analysis sample consists of recently displaced (displaced in last 5 years) and non-migrants who were born in the survey country of the indicated gender and age at survey. Data is pooled over all countries, excluding Pakistan (since partition occurred 26 years prior to survey year) and India (as none of the Indian surveys include birthplace information).



(a) Women

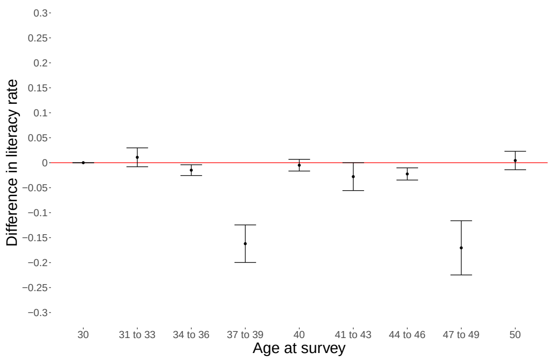


(b) Men

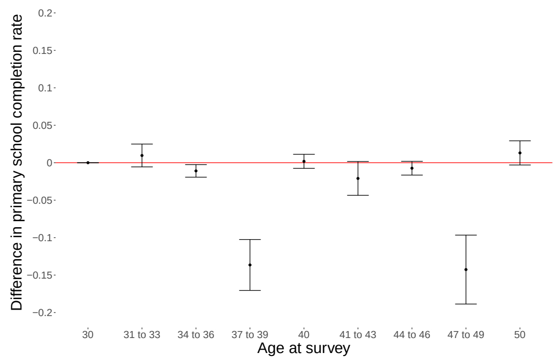
Figure 2.A.3: Cross-country evidence: Effect of displacement on marriage rates (clustered standard errors)

Notes: 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 omitted age group is age 38 at time of survey. 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. The analysis sample consists of recently displaced (displaced in last 5 years) and non-migrants of the indicated gender and age at survey. Data is pooled over all countries, excluding Pakistan (since partition occurred 26 years prior to survey year).

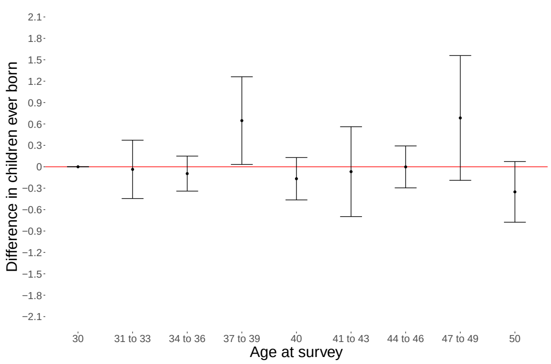
Appendix B: Robustness Checks for Pakistan Evidence



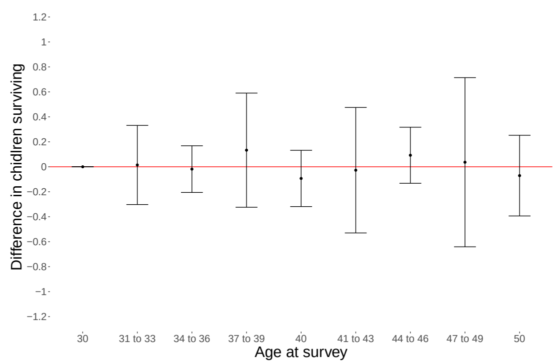
(a) Literacy



(b) Completed primary school



(c) Children every born



(d) Children surviving

Figure 2.B.1: Pakistan evidence: Parallel trends

Notes: This figure plots parallel trends tests for additional 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. Individual observations are weighted by inverse sampling probability. The dependent variable "Child marriage" is an indicator whether the respondent was married before the age of 18. The dependent variable "Literate" is an indicator of whether the respondent is literate (regardless of formal school attendance). The dependent variable "Completed primary school" is an indicator of whether the respondent completed primary school. 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 survey.

Table 2.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 at partition
	(1)	(2)	(3)	(4)
Young X Displaced	.01***	.038	-.28	.062***
	(.002)	(.044)	(.317)	(.023)
Young	-.013***	.06**	-.63***	.23***
	(.002)	(.029)	(.236)	(.017)
Displaced	.00024	.019	-.25	-.015***
	(.001)	(.034)	(.261)	(.005)
Comparison mean	.997	.308	19.3	.0231
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 at partition
	(1)	(2)	(3)	(4)
Young X Displaced	-.014*	-.0043	.14	.015***
	(.008)	(.043)	(.305)	(.005)
Young	-.042***	.13***	-1.2***	-.023***
	(.005)	(.026)	(.214)	(.004)
Displaced	.00024	.019	-.25	-.015***
	(.001)	(.034)	(.261)	(.005)
Comparison mean	.997	.308	19.3	.0231
Adjusted R^2	.01	.01	.02	.02
N	35209	32967	32967	32967

Notes: Each column represents a single regression estimated from Equation 2.2. The data are from the 1973 Pakistan HED survey. 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. "Displaced" is an indicator of whether the respondent was displaced from India around the time of partition (between 1947-1949). 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 at partition" is an indicator variable of whether a respondent was first married between 1946 - 1949. 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 2.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	2.4e-16	.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

Notes: Each column represents a single regression estimated from Equation 2.2. The data are from the 1973 Pakistan HED survey. 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. "Young" is an indicator of whether the respondent was between the ages of 10 and 17 at the time of partition. "Displaced" is an indicator of whether the respondent was displaced from India at the time of partition. 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 survey. 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 2.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

Notes: Each column represents a single regression estimated from Equation 2.2. The data are from the 1973 Pakistan HED survey. 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. "Young" is an indicator of whether the respondent was between the ages of 10 and 17 at the time of partition. "Displaced" is an indicator of whether the respondent was displaced from India 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: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Chapter 3

Cash Transfers and Voting Behavior

3.1 Introduction

Cash transfer programs are a popular tool used in many developing countries to fight poverty and provide social protection. Even though certain program characteristics vary from country to country, in general cash transfer programs provide money to poor households either conditional on human investment - conditional cash transfers (CCTs), or unconditionally. These have been subject to many studies and evaluations, and have an evidence base that is arguably one of the highest among poverty reduction tools. Evaluations of cash transfer programs in parts of Africa, Asia and Latin America have shown that such programs have a wide variety of impacts on socio-economic outcomes such as improvements in child health (Duflo (2003), Cunha et al. (2010)), increase in child schooling (Edmonds (2006), Baird et al. (2011)), and improvements in household income (Sadoulet et al. (2001), Attanasio and Mesnard (2006)).

There are political economy aspects of such programs as well. Theory tells us that anti-poverty programs can influence individual political participation and voting preferences either through voter-driven channels - recipients value transfers and reward incumbents, or infer competence of politicians through social policy choices; or through policy maker-driven channels

- politicians can strategically allocate spending to benefit certain groups to raise support in their favor (Drazen and Eslava, 2010). Regardless of the mechanism, there is plenty of evidence that shows that cash transfer programs (conditional or otherwise) affects voting behavior. Baez et al. (2012) show that *Familias en Accion*, a large-scale anti-poverty program in Colombia, increases voter registration and voting in favor of the implementing party. Manacorda et al. (2011) use self-reported data to find that beneficiaries of an unconditional cash transfer program in Uruguay (PANES) express larger support for the incumbent that implemented the program. De La O (2013) and Zucco Jr (2013) show similar findings for the *Oportunidades* CCT program in Mexico and *Bolsa Familia* CCT program in Brazil respectively.

I add to this literature by providing supporting evidence for the largest unconditional cash transfer program in Pakistan - the Benazir Income Support Program (BISP). Using privately collected survey data and matching it with administrative data from the program, I try to study the impacts of BISP on voting behavior by using a regression discontinuity (RD) design that leverages the eligibility threshold in the program's proxy means test targeting approach. However, I revert to partial identification using bounds after finding significant manipulation in the running variable. In the absence of informative bounds, I then provide biased RD estimates to show correlations between voting outcomes of interest and being a BISP beneficiary.

I find that the program is correlated with an increased likelihood of voting. I also find that being a beneficiary is correlated with an increased likelihood of voting for the incumbent political party that initiated the program relative to non-beneficiaries, which does not come at a cost to the main opposition. These results are in line with what De La O (2013) find in Mexico, and suggest that the bonus for the incumbent party can be explained by higher voter base rather than turn over.

3.2 Background

3.2.1 Political Context

Pakistan is a federal state with a multi-party political system. Elections are held every 5 years to elect members of provincial and national assemblies that constitute the provincial and federal governments respectively. While there are many political parties and party factions, there are two that stand out as most popular: Pakistan Muslim League (N) (PML-N), and Pakistan People's Party (PPP). Although they subscribe to different ideologies, both parties share important similarities. Both have held power multiple times during the country's short history; both practice some form of dynastic politics; both have provincial strongholds (south eastern province of Sindh for PPP (Das, 2001) and north eastern province of Punjab for PML Shah and Asif (2015)); and both have central lead figures who drive party popularity.

From the mid-1980's to mid-2000's, this central figure for PPP was Benazir Bhutto who not only led the party during this time, but was also elected Prime Minister in the 1988 and 1993 general elections. However, she was assassinated in December 2007 while campaigning for the 2008 general elections. Despite this setback, PPP went on to secure the majority vote and formed the ruling coalition. Their term in office was the first in the history of the country to complete a 5 year term, with general elections held again in 2013.

3.2.2 Benazir Income Support Program (BISP)

Benazir Income Support Program (BISP), is a large-scale social protection program that was first initiated on a small scale in 2008 and then extended nationwide in 2011. The program was named in honor of Benazir Bhutto, the recently deceased ex-Prime Minister and party leader of the recently elected Pakistan People's Party (PPP). Unlike most large-scale programs around the world, the program name clearly signals the party responsible for it and remains unchanged even after 8 years and two new regimes.

The program initially began as an unconditional cash transfer program but now provides welfare support in the form of vocational training and educational support for qualifying low-income families. As of now the program has an annual budget of $1.15 billion, has 5.4 million beneficiaries, and accounts for nearly 0.3% of the GDP of Pakistan. The cash transfer program provides support exclusively to a female member of eligible household through quarterly cash payments of Rs. 4,834 (approximately $32). Eligibility for the program is determined by a Proxy Means Test (PMT) score.

In order to construct a poverty score BISP conducted nationwide household surveys through different non-government organizations and survey firms which collected information on household demographics and asset ownership. The PMT score is calculated with an algorithm that weighs several variables to predict household income and welfare. The score takes a value from 0 - 100 with lower values indicating higher poverty. A household is eligible for the transfer if their PMT score is below 16.17, there is an adult female in the household, and the adult female has a Computerized National Identification Card (CNIC)[1]. CNICs are a universal government identification in Pakistan and are a prerequisite in order to obtain a driver's license or to be able to vote.

3.3 Data

This paper uses data from two different sources to study the effect of BISP on voting behavior: (1) Household survey data for a random sample of rural households in 3 districts of southern Punjab, and (2) administrative data from BISP. Details on these data and how they are matched are given below.

The household survey data was collected in November 2013 for a separate study[2]. These

[1]A Computerized National Identity Card (CNIC) provides a unique identification for Pakistani Citizens similar to *Aadhaar* cards in India. A CNIC is required to obtain formal public services and to vote.

[2]Household survey data was collected by Cheema et al. (2020) which I'm using with the authors' permission.

survey collected data on household demographics, CNIC numbers of female members, payments received from BISP, and voting behavior of the female head of the household.

I use CNIC information collected in the household survey to match women in the survey sample with administrative data from BISP. This data contains households' PMT score, their survey interview date, and total payments made to the household until October 2017.

The timeline of different data collection and BISP related events for this sample is given below:

Apr 2011 - Nov 2011: Survey conducted by BISP for PMT score calculation

Oct 2011 - to date: Eligible households made cash transfer payments through BISP

Nov 2013 - Dec 2013: Independent household survey

The matching exercise between my two data sources yields a sample of nearly 6,200 households for whom I have both survey data and BISP data. However, close examination of BISP administrative data for this sample reveals problems in implementation of the program in the form of payments not made to eligible households, and payments made to ineligible households. Nearly 37% of households with a PMT score below 16.17 (i.e. eligible households) are not made a payment from BISP, and 3% of households with a PMT score of above 16.17 (i.e. ineligible households) are made a payment from BISP. Table 3.1 provides a breakdown of BISP eligible and ineligible households that are made a payment from the program.

Furthermore, data on disbursement of BISP payments (from administrative data) does not match data on receipt of these funds by beneficiaries (from survey data). According the household survey data, 86% of eligible households and almost 6.5% ineligible households received payments from BISP (see table 3.2). Ignoring the reason for this mismatch, I treat the latter as the first stage of the program.

3.4 Estimation

Under normal circumstances, I would identify the causal impact of BISP on voting behavior by using a Regression Discontinuity (RD) framework and exploiting the discontinuous nature of program eligibility requirements. The standard RD estimation in this case would be given by:

$$Y = \alpha_0 + \beta_1 \mathbf{1}(PMT \leq 0) + f_l(PMT) + g_r(PMT) \tag{3.1}$$

Where Y are outcomes of interest, PMT is the running variable rescaled such that the cutoff for program eligibility is now at 0, and f_l and g_r are functions of the running variable to the left and right of the threshold respectively.

Circumstances for this context and analysis sample, however, are not normal. Validity of a standard RD design requires treatment assignment at the threshold to be as good as random; or, stated another way, the distribution of units' potential outcomes needs to vary continuously with the running variable around the threshold. Testing this in my context using the McCrary (2008) test of manipulation shows a large and significant difference in the density of the running variable to the left and right of the cutoff. This renders the use of standard RD in this context invalid.

3.4.1 Identification with Manipulated Running Variable

Given the manipulation in the running variable as indicated by McCrary (2008), I consider two non-standard RD methods to identify the effects of BISP on voting outcomes: "Donut-hole" RD and bounds on the RD estimate.

Donut-hole RD

Donut-hole RD is an RD design in which observations closest to the threshold where sorting is present are omitted from the analysis. While this design has been used in the literature (Almond and Doyle, 2011; Barreca et al., 2011), validity of its use is still somewhat debated

(Eggers et al., 2015). Regardless of this designs criticisms, it is not feasible in my context because the "donut-hole" needed to account for heaps in data would remove a relatively large portion of the running variable around the threshold. The histogram in figure 3.2 shows heaps as far as 2 PMT point to the left of the cut-off. I therefore turn to estimating bounds of the treatment effect.

Bounds

Many applied papers have used versions of bounding techniques established by Lee (2009) in RD settings (Card et al., 2009; Anderson and Magruder, 2012). However, these only account for sharp RDs. I use the approach and estimation procedure developed in Gerard et al. (2020) to determine bounds on the treatment effect of a fuzzy RD.

The framework developed in Gerard et al. (2020) assumes that units are either *potentially-assigned*, for whom standard assumption of RD designs holds, or are *always-assigned*, for whom the running variable takes a value only to one side of the cutoff. They calculate bounds of treatment effects for potentially-assigned units by fist estimating the proportion of always-assigned units using the magnitude of the discontinuity in density of the running variable, then bound treatment effects by finding the "highest" and "lowest" value of outcomes for always assigned units. For sharp RDs, bounds are obtained by trimming the tails of the outcome distribution among units close to the threshold in the manipulated side. Construction of bounds in the fuzzy RD case are more elaborate, details for which can be found in Gerard et al. (2020).

Estimating these bounds assumes (in addition to the standard RD assumptions) one-sided manipulation, i.e. realization of the running variable for manipulated units is always on one side of the cutoff. In this setting manipulation would be to the left of the cutoff (i.e. towards eligibility) and there's sufficient reason to assume that it would be one-sided because: (i) if manipulation is coming from households misreporting assets, they would do so to qualify for the program; and (ii) if manipulation is politically driven, it would be to have more beneficiaries to have a higher political return.

3.5 Results

Figure 3.3 graphically illustrates the first stage. Being eligible for the program (i.e. being below the threshold) increases the probability of receiving BISP payments by 70% pt. As mentioned before, this first stage comes from receipt of BISP payments as reported in the household survey data. This differs from being paid by BISP, the data for which comes from administrative data. Also note that this first stage is for the entire sample and does not take into account always-assigned (or manipulated) units.

3.5.1 Bounds

Results from the bounding exercise are given in table 3.3 using a bandwidth of 8 around the cutoff. Estimates in this table are calculated using Stata code provided by Gerard et al. (2020).

Panel 3.3a gives estimates of the inputs that go into calculating bounds. The share of always-assigned individuals is estimated to be 0.293, which is very similar to the discontinuity in the density of the running variable as estimated by McCrary (2008) test of manipulation (figure 3.1). Taking this into account, the increase in take up at the cut-off for potentially assigned units is estimated at almost 52%. This is the effective first stage of the RD.

Panels 3.3b, 3.3c and 3.3d give bounds on the intent-to-treat (ITT - sharp RD) and local average treatment effect (LATE- fuzzy RD) of BISP on turnout, voting for PPP - the incumbent, and voting for PML - the historically popular party in this sample respectively. Given the high proportion of always-assigned units, the bounds are quite uninformative. Bounds on both estimates for each outcome variable contain 0. What stands out from these results, however, is that bounds for the outcome variable "voted for PML" are centered around 0, whereas bounds for the other two outcome variables have a larger positive value as compared to a negative one. This provides suggestive evidence that BISP may have increased voter turn out and increased the likelihood of voting for PPP.

For completeness, I estimate these bounds using bandwidths of 6 and 10 which yields similar results. These are given in the Appendix 3.6.

3.5.2 Correlations

In the absence of informative results from the bounding exercise above, I revert to providing correlational evidence using biased coefficients from estimating equation 3.1 while including the manipulated sample. This estimation is graphically illustrated in figures 3.4, 3.5 and 3.6, and coefficients are given in table 3.4.

Biased estimates from estimating a sharp RD while including manipulated units suggests a statistically significant correlation between being a BISP beneficiary and higher voter turnout, and voting for PPP (almost 10% point).

3.6 Conclusion

Although lacking clear causal interpretation, my results suggest that being eligible for the cash transfer program has positive effects for the incumbent political party that introduced the program. However, higher vote share for the incumbent party does not seem to come at the cost of lower vote share for the main opposition and historically popular party. It seems that higher votes for the incumbent come from increased voter turn out. These findings are in line with findings in existing literature on the effect of large scale welfare program on voting outcomes in other parts of the world.

Figures and Tables

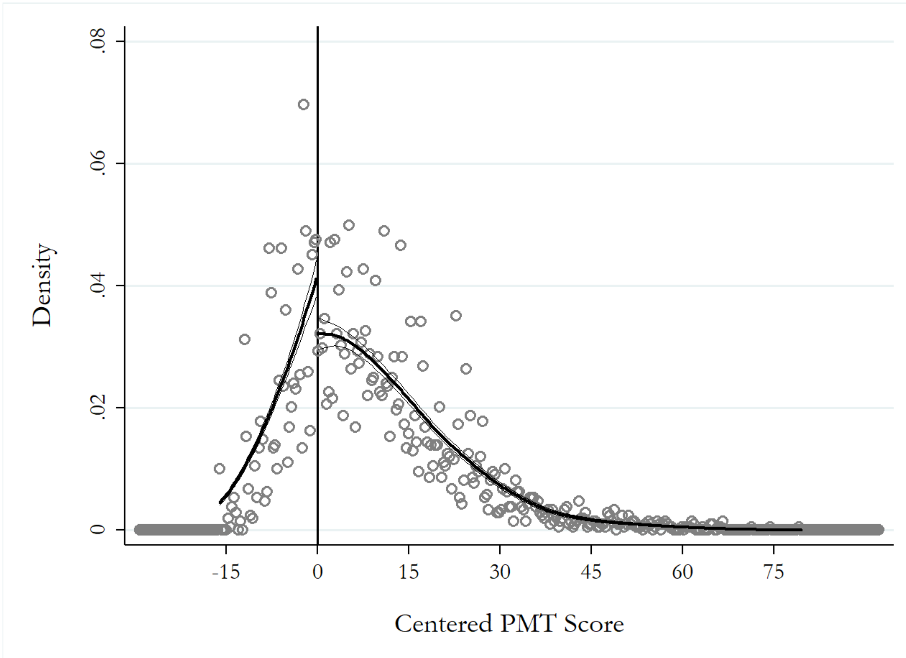


Figure 3.1: RD test for manipulation

Notes: This graph implements the McCrary (2008) test for manipulation. Data are household level administrative data from BISP for a representative sample of households in three southern districts of Punjab, Pakistan. Each circle is the density of observations within a bin size of 0.34 of the centered PMT score. Dark solid lines are kernel densities to the left and right of the cutoff, and light solid lines are associated confidence intervals. The test gives a significant absolute discontinuity estimate of 0.26.

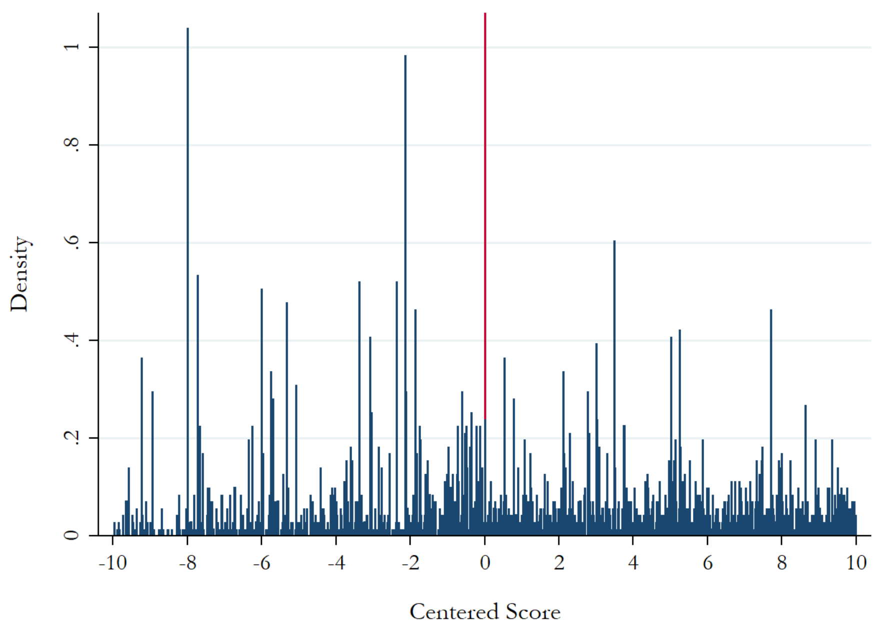


Figure 3.2: Histogram of the running variable

Notes: This graph plots a histogram showing the density of the running variable in bin sizes of 0.02. Data are household level administrative data from BISP for a representative sample of households in three southern districts of Punjab, Pakistan. The red line represents the threshold for program eligibility.

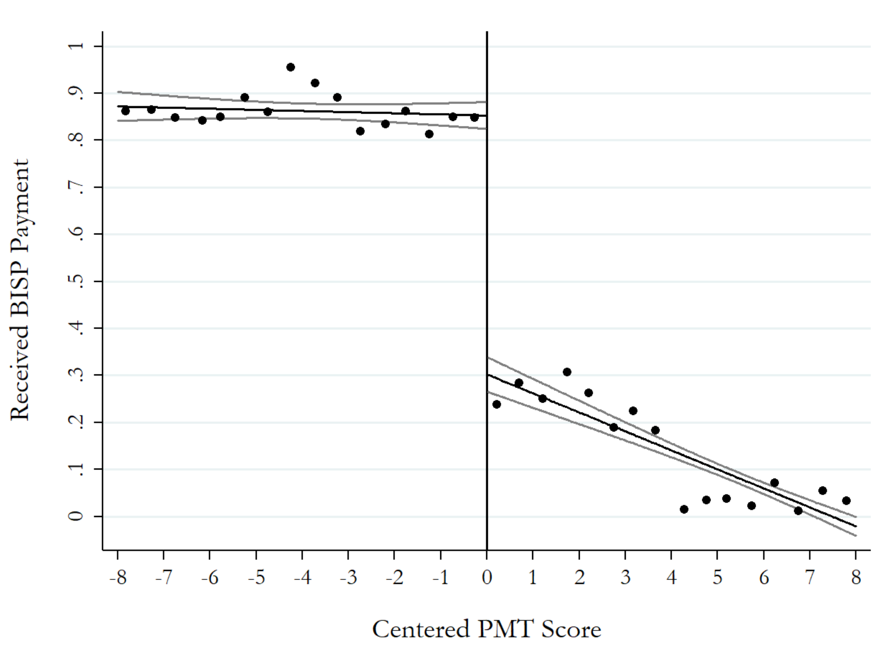


Figure 3.3: First stage of RD: Probability of receiving BISP

Notes: This graph gives the first stage of the RD. "Received BISP payment" is a binary variable which indicates if a household received any payment from BISP. Data is survey data at the household level and is from a representative sample of households in three southern districts of Punjab, Pakistan. Each dot gives the mean over a bin of 0.5 of the centered PMT score. Dark solid lines are predicted values based on first-order polynomial regressions, and light solid lines are confidence intervals of the prediction.

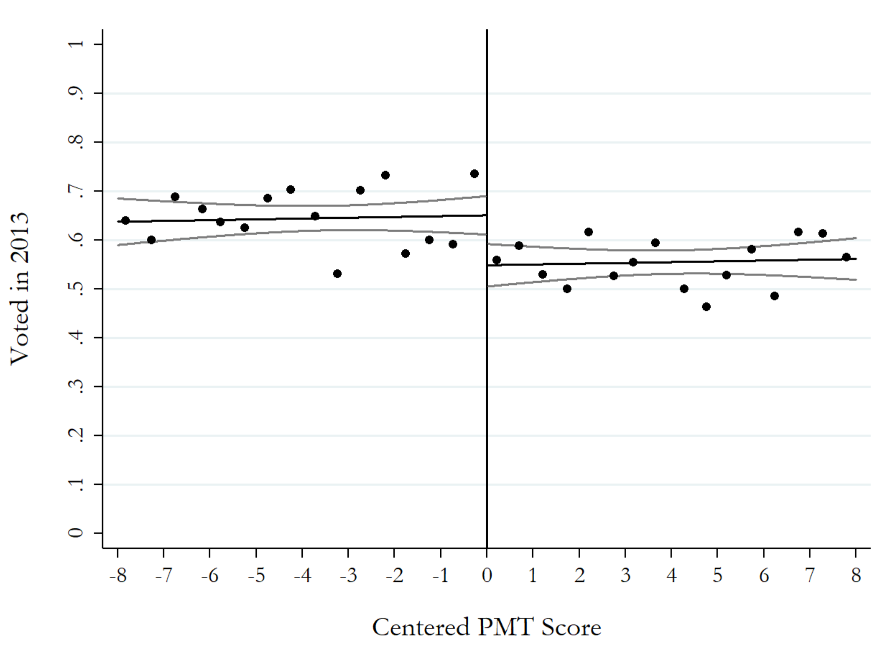


Figure 3.4: Correlational evidence for voter turnout in 2013 general elections

Notes: This graphs provides correlational evidence because it includes the manipulated sample. The outcome is a binary variable which indicates if household head voted in the 2013 general elections. Data is survey data at the household level and is from a representative sample of households in three southern districts of Punjab, Pakistan. Each dot gives the mean over a bin of 0.5 of the centered PMT score. Dark solid lines are predicted values based on first-order polynomial regressions, and light solid lines are confidence intervals of the prediction.

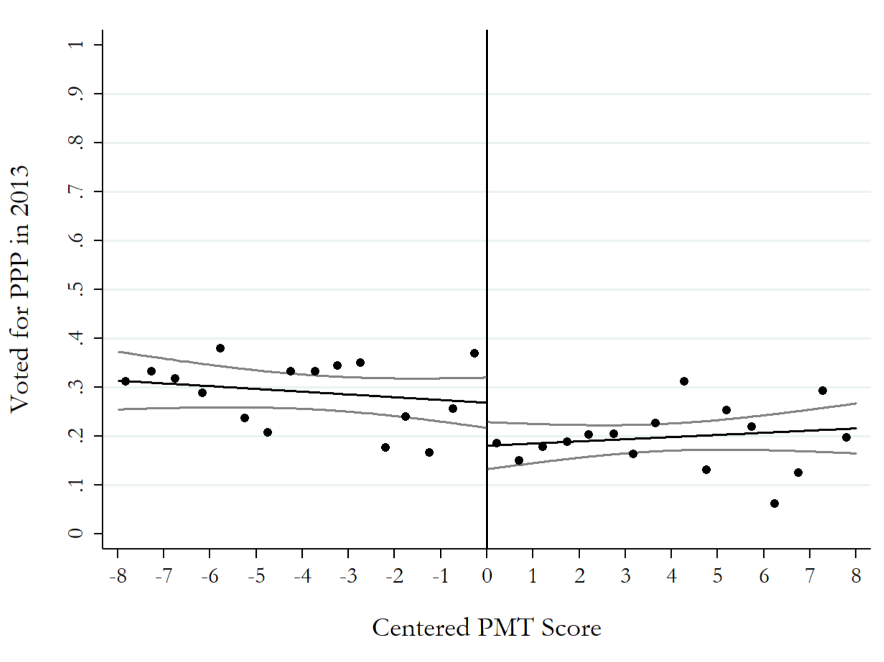


Figure 3.5: Correlational evidence for voting for PPP in 2013 general elections

Notes: This graph provides correlational evidence because it includes the manipulated sample. The outcome is a binary variable which indicates if household head voted for PPP (the incumbent political party) in the 2013 general elections. Data is survey data at the household level and is from a representative sample of households in three southern districts of Punjab, Pakistan. Each dot gives the mean over a bin of 0.5 of the centered PMT score. Dark solid lines are predicted values based on first-order polynomial regressions, and light solid lines are confidence intervals of the prediction.

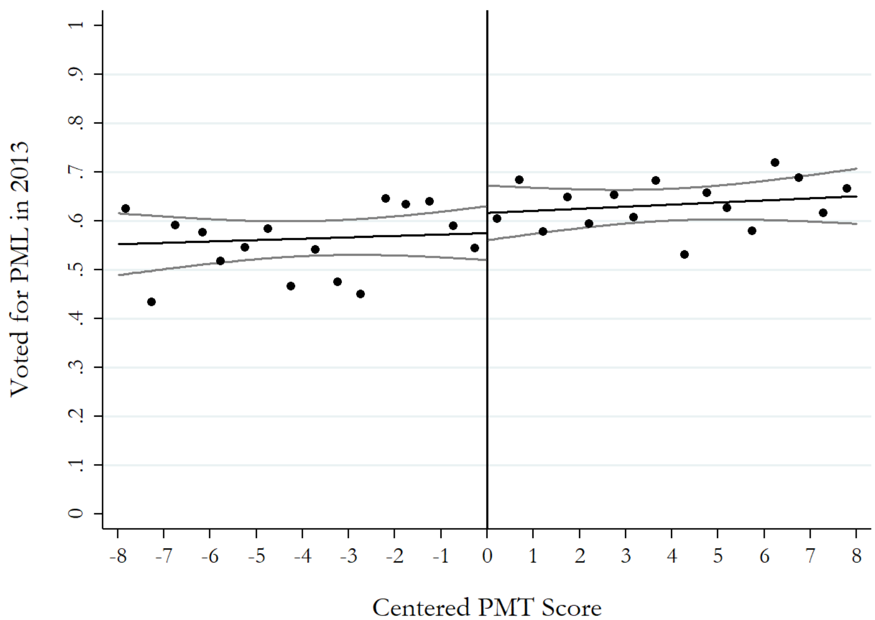


Figure 3.6: Correlational evidence for voting for PML in 2013 general elections

Notes: This graph provides correlational evidence because it includes the manipulated sample. The outcome is a binary variable which indicates if household head voted for PML (the historically popular political party in Punjab, Pakistan) in the 2013 general elections. Data is survey data at the household level and is from a representative sample of households in three southern districts of Punjab, Pakistan. Each dot gives the mean over a bin of 0.5 of the centered PMT score. Dark solid lines are predicted values based on first-order polynomial regressions, and light solid lines are confidence intervals of the prediction.

Table 3.1: BISP payments made by program eligibility

	Eligible		Ineligible	
	Mean	Mean	Mean	Mean
Paid BISP payments	.628		.0309	
	(.0113)		(.0026)	
Total BISP payments if paid		89757		75645
		(462.57)		(1425.74)
Observations	1846	1160	4332	134

Notes: This table highlights issues with implementation of BISP. Data are household level administrative data from BISP for a representative sample of households in three southern districts of Punjab, Pakistan. "Eligible" refers to households to the left of the cutoff that are eligible to receive BISP payments. "Ineligible" refers to households to the right of the cutoff that are not eligible to receive BISP payments. "Paid BISP payments" is a binary variable indicating if BISP made payments to the household (as opposed to *households receiving BISP payments* which is used to show the first stage), and "Total BISP payments if paid" is the average of total BISP payments (in Rs.) made to households conditional on being paid.

Table 3.2: Probability of receiving BISP

	Eligible	Ineligible
Received BISP payments	.865	.0646
	(.00795)	(.00374)
Observations	1846	4332

Notes: Data are household level survey data for a representative sample of households in three southern districts of Punjab, Pakistan. "Eligible" refers to households to the left of the cutoff that are eligible to receive BISP payments. "Ineligible" refers to households to the right of the cutoff that are not eligible to receive BISP payments. "Received BISP payments" is a binary variable indicating if households received any payment from BISP. This graph does not give the first stage of the RD as it gives averages for whole sample.

Table 3.3: Estimated treatment effects of BISP on voting behavior in 2013 general elections

(a) Basic inputs

	Estimate	95% CI
Share of always-assigned individuals	0.293	[0.184; 0.401]
Increase in takeup at cutoff	0.525	[0.465; 0.586]

(b) Voter turnout

	Bounds	95% CI
ITT/ Sharp RD	[-0.049; 0.365]	[-0.143; 0.487]
LATE/ Fuzzy RD	[-0.128; 0.561]	[-0.35; 0.69]

(c) Voted for PPP

	Bounds	95% CI
ITT/ Sharp RD	[-0.174; 0.291]	[-0.227; 0.413]
LATE/ Fuzzy RD	[-0.202; 0.676]	[-0.284; 0.961]

(d) Voted for PML

	Bounds	95% CI
ITT/ Sharp RD	[-0.299; 0.347]	[-0.451; 0.472]
LATE/ Fuzzy RD	[-0.633; 0.501]	[-0.84; 0.651]

Notes: Total number of observations within bandwidth of 8 around the cutoff is 3,097 BISP eligible and ineligible households. Confidence intervals are based on 1,000 bootstrap samples. Data are household level survey data from three southern districts of Punjab, Pakistan. Outcomes in panels (b), (c), and (d) are binary variables which indicates if household head voted, if household head voted for PPP, and if household head voted for PML in the 2013 general elections respectively.

Table 3.4: Correlational evidence on the effect of BISP on voting behavior in 2013 general elections

(a) Voter turnout

	Estimate	95% CI
Sharp RD	0.094	[0.019; 0.169]
Fuzzy RD	0.159	[0.01; 0.308]

(b) Voted for PPP

	Estimate	95% CI
Sharp RD	0.109	[0.027; 0.19]
Fuzzy RD	0.218	[0.051; 0.385]

(c) Voted for PML

	Estimate	95% CI
Sharp RD	-0.033	[-0.129; 0.062]
Fuzzy RD	-0.088	[-0.283; 0.107]

Notes: Total number of observations within bandwidth of 8 around the cutoff is 3,097 BISP eligible and ineligible households. Confidence intervals are based on 1,000 bootstrap samples. Estimates are not causal because data includes and does not account for manipulated sample. Data are household level survey data from three southern districts of Punjab, Pakistan. Outcomes in panels (a), (b), and (c) are binary variables which indicates if household head voted, if household head voted for PPP, and if household head voted for PML in the 2013 general elections respectively.

Appendix: Robustness Checks

Table 3.A.1: Estimated treatment effects of BISP on voting behavior in 2013 general elections (bandwidth = 6)

(a) Basic inputs

	Estimate	95% CI
Share of always-assigned individuals	0.346	[0.235; 0.457]
Increase in takeup at cutoff	0.524	[0.453; 0.595]

(b) Voter turnout

	Bounds	95% CI
ITT/ Sharp RD	[-0.087; 0.431]	[-0.198; 0.528]
LATE/ Fuzzy RD	[-0.215; 0.571]	[-0.487; 0.681]

(c) Voted for PPP

	Bounds	95% CI
ITT/ Sharp RD	[-0.164; 0.36]	[-0.221; 0.515]
LATE/ Fuzzy RD	[-0.191; 0.869]	[-0.279; 1.143]

(d) Voted for PML

	Bounds	95% CI
ITT/ Sharp RD	[-0.351; 0.373]	[-0.534; 0.489]
LATE/ Fuzzy RD	[-0.685; 0.506]	[-0.916; 0.658]

Notes: Total number of observations within bandwidth of 6 around the cutoff is 2,376 BISP eligible and ineligible households. Confidence intervals are based on 1,000 bootstrap samples. Data are household level survey data from three southern districts of Punjab, Pakistan. Outcomes (panels (b), (c), and (d)) are binary variables which indicates if household head voted, if household head voted for PPP, and if household head voted for PML in the 2013 general elections respectively.

Table 3.A.2: Estimated treatment effects of BISP on voting behavior in 2013 general elections (bandwidth = 10)

(a) Basic inputs

	Estimate	95% CI
Share of always-assigned individuals	0.254	[0.158; 0.349]
Increase in takeup at cutoff	0.542	[0.486; 0.599]

(b) Voter turnout

	Bounds	95% CI
ITT/ Sharp RD	[-0.02; 0.32]	[-0.103; 0.429]
LATE/ Fuzzy RD	[-0.058; 0.544]	[-0.236; 0.673]

(c) Voted for PPP

	Bounds	95% CI
ITT/ Sharp RD	[-0.177; 0.257]	[-0.227; 0.359]
LATE/ Fuzzy RD	[-0.202; 0.565]	[-0.271; 0.816]

(d) Voted for PML

	Bounds	95% CI
ITT/ Sharp RD	[-0.279; 0.291]	[-0.411; 0.426]
LATE/ Fuzzy RD	[-0.61; 0.485]	[-0.801; 0.656]

Notes: Total number of observations within bandwidth of 10 around the cutoff is 3,556 BISP eligible and ineligible households. Confidence intervals are based on 1,000 bootstrap samples. Data are household level survey data from three southern districts of Punjab, Pakistan. Outcomes (panels (b), (c), and (d)) are binary variables which indicates if household head voted, if household head voted for PPP, and if household head voted for PML in the 2013 general elections respectively.

Bibliography

Almond, Douglas and Joseph J Doyle, "After midnight: A regression discontinuity design in length of postpartum hospital stays," *American Economic Journal: Economic Policy*, 2011, *3* (3), 1–34.

Anderson, Michael and Jeremy Magruder, "Learning from the crowd: Regression discontinuity estimates of the effects of an online review database," *The Economic Journal*, 2012, *122* (563), 957–989.

Attanasio, Orazio and Alice Mesnard, "The impact of a conditional cash transfer programme on consumption in Colombia," *Fiscal studies*, 2006, *27* (4), 421–442.

Atuesta, Laura H, Oscar S Siordia, and Alejandro Madrazo Lajous, "The "war on drugs" in Mexico:(official) database of events between December 2006 and November 2011," *Journal of conflict resolution*, 2019, *63* (7), 1765–1789.

Baez, Javier E, Adriana Camacho, Emily Conover, and Román A Zárate, *Conditional cash transfers, political participation, and voting behavior*, The World Bank, 2012.

Baird, Sarah, Craig McIntosh, and Berk Özler, "Cash or condition? Evidence from a cash transfer experiment," *The Quarterly Journal of Economics*, 2011, p. qjr032.

Barreca, Alan I, Melanie Guldi, Jason M Lindo, and Glen R Waddell, "Saving babies? Revisiting the effect of very low birth weight classification," *The Quarterly Journal of Economics*, 2011, *126* (4), 2117–2123.

Becker, Gary S, "Crime and punishment: An economic approach," in "The economic dimensions of crime," Springer, 1968, pp. 13–68.

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

_ , Irena Grosfeld, Pauline Grosjean, Nico Voigtländer, and Ekaterina Zhuravskaya, "Forced migration and human capital: evidence from post-WWII population transfers," *American Economic Review*, forthcoming.

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

— **and Rinchan Ali Mirza**, "Displacement and development: Long term impacts of population transfer in India," *Explorations in Economic History*, 2019, *73*.

— **, Asim Khwaja, and Atif Mian**, "The big march: migratory flows after the partition of India," *Economic and Political Weekly*, 2008, pp. 39–49.

Bradley, Megan, *Refugee repatriation: justice, responsibility and redress*, Cambridge University Press, 2013.

Brookings Institution, "Improving the Protection of Internally Displaced Women: Assessment of Progress and Challenges," Technical Report 2014.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

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

Canay, Ivan A, Andres Santos, and Azeem M Shaikh, "The Wild Bootstrap with a "Small" Number of "Large" Clusters," *The Review of Economics and Statistics*, Forthcoming.

Card, David, Carlos Dobkin, and Nicole Maestas, "Does Medicare save lives?," *The quarterly journal of economics*, 2009, *124* (2), 597–636.

Carrell, Scott E, Teny Maghakian, and James E West, "A's from Zzzz's? The causal effect of school start time on the academic achievement of adolescents," *American Economic Journal: Economic Policy*, 2011, *3* (3), 62–81.

Cheema, Ali, Asim I Khwaja, Farooq Naseer, and Jacob N Shapiro, "Glass Walls: Experimental Evidence on Access Constraints Faced by Women," 2020.

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

Corno, Lucia, Nicole Hildebrandt, and Alessandra Voena, "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments," *Econometrica*, Forthcoming.

Cunha, Jesse M et al., "Testing Paternalism: Cash vs. In-kind Transfers in Rural Mexico," *Unpublished manuscript*, 2010.

Das, Suranjan, *Kashmir and Sindh: Nation-building, ethnicity and regional politics in South Asia*, Anthem Press, 2001.

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

Dell, Melissa, "Trafficking networks and the Mexican drug war," *American Economic Review*, 2015, *105* (6), 1738–79.

Delprato, Marcos, Kwame Akyeampong, and Máiréad Dunne, "Intergenerational education effects of early marriage in Sub-Saharan Africa," *World Development*, 2017, *91*, 173–192.

Doleac, Jennifer L and Nicholas J Sanders, "Under the cover of darkness: How ambient light influences criminal activity," *Review of Economics and Statistics*, 2015, *97* (5), 1093–1103.

Drazen, Allan and Marcela Eslava, "Electoral manipulation via voter-friendly spending: Theory and evidence," *Journal of development economics*, 2010, *92* (1), 39–52.

Duflo, Esther, "Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa," *The World Bank Economic Review*, 2003, *17* (1), 1–25.

Edmonds, Eric V, "Child labor and schooling responses to anticipated income in South Africa," *Journal of development Economics*, 2006, *81* (2), 386–414.

Eggers, Andrew C, Anthony Fowler, Jens Hainmueller, Andrew B Hall, and James M Snyder Jr, "On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races," *American Journal of Political Science*, 2015, *59* (1), 259–274.

Evalúa, México, "Seguridad y justicia penal en los estados: 25 indicadores de nuestra debilidad institucional," *Recuperado de http://www. mexicoevalua. org/wp-content/uploads/2013/02/MEX-EVA_INDX_SJPE-LOW. pdf*, 2012.

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

Frattini, Tommaso, Francesco Fasani, and Luigi Minale, "(The Struggle for) Refugee Integration into the Labour Market: Evidence from Europe," 2020.

Gerard, François, Miikka Rokkanen, and Christoph Rothe, "Bounds on treatment effects in regression discontinuity designs with a manipulated running variable," *Quantitative Economics*, 2020, *11* (3), 839–870.

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

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

IEP, "Mexico peace index 2013," Technical Report 2013.

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

Jagnani, Maulik, "Poor sleep: Sunset time and human capital production," 2018.

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

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

Jensen, Robert and Rebecca Thornton, "The Consequences of Early Marriage in the Developing World," *Oxfam Journal of Gender and Development*, 2003, *11* (2), 9–19.

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

Jr, Cesar Zucco, "When payouts pay off: Conditional cash transfers and voting behavior in Brazil 2002–10," *American journal of political science*, 2013, *57* (4), 810–822.

Keygnaert, Ines, 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*, 2012, *14* (5), 505–520.

Khalidi, Omar, "From Torrent to Trickle: Indian Muslim Migration to Pakistan, 1947—97," *Islamic studies*, 1998, *37* (3), 339–352.

Lee, David S, "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *The Review of Economic Studies*, 2009, *76* (3), 1071–1102.

Maia, Maia, Nasma Berri, and Sawsan Abdulrahim, "Marriage and fertility patterns among Jordanians and Syrian refugees in Jordan," in "in" Economic Research Forum 2018.

Manacorda, Marco, Edward Miguel, and Andrea Vigorito, "Government transfers and political support," *American Economic Journal: Applied Economics*, 2011, *3* (3), 1–28.

Masterson, Amelia Reese, 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*, 2014, *14* (1), 25.

McCollister, Kathryn E, Michael T French, and Hai Fang, "The cost of crime to society: New crime-specific estimates for policy and program evaluation," *Drug and alcohol dependence*, 2010, *108* (1-2), 98–109.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of econometrics*, 2008, *142* (2), 698–714.

Meeus, Jean H, *Astronomical algorithms*, Willmann-Bell, Incorporated, 1991.

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

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

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

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

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

O, Ana L De La, "Do conditional cash transfers affect electoral behavior? Evidence from a randomized experiment in Mexico," *American Journal of Political Science*, 2013, *57* (1), 1–14.

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

Raj, Anita, 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*, 2019, *19* (1), 1–20.

Sadoulet, Elisabeth, Alain De Janvry, and Benjamin Davis, "Cash transfer programs with income multipliers: PROCAMPO in Mexico," *World development*, 2001, *29* (6), 1043–1056.

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

Shah, Aqil and Bushra Asif, "Pakistan in 2014: Democracy under the Military's Shadow," *Asian Survey*, 2015, *55* (1), 48–59.

Sieverding, Maia, 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," Vol. 1257 2018.

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

Smith, Austin C, "Spring forward at your own risk: Daylight saving time and fatal vehicle crashes," *American Economic Journal: Applied Economics*, 2016, *8* (2), 65–91.

Steele, Abbey et al., "Massive Civilian Displacement in Civil War: Assessing Variation in Colombia," Technical Report, Households in Conflict Network 2007.

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

UNHCR, "Global Trends: Forced Displacement in 2018," Technical Report 2019.

UNICEF et al., "Sexual and gender-based violence against refugees, returnees and internally displaced persons-guidelines for prevention and response," 2003.

UNODC, *Global study on homicide 2013: trends, contexts, data* 2013.

Usta, Jinan and Amelia Reese Masterson, "Women and health in refugee settings: The case of displaced Syrian women in Lebanon," in "Gender-based violence," Springer, 2015, pp. 119–143.

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

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

Vu, Alexander, 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*, 2014, *6.*

Williams, Phil, "The terrorism debate over Mexican drug trafficking violence," *Terrorism and Political Violence*, 2012, *24* (2), 259–278.

Wirtz, Andrea L, 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*, 2014, *8* (1), 10.