Schooling and Marriage Markets
How an Education Reform Impacted Matching Patterns:
Evidence from Zimbabwe

Isabelle Salcher
New York University
January 31, 2020

Abstract

This paper examines the impact of an aggregate shock to human capital on the marriage market. In Zimbabwe, a large nationwide education reform together with a large age difference between spouses allows us to isolate the impact an increase in female education has on matching patterns. Exploiting the age-specific exposure to the reform in a regression discontinuity design, I find that females who were still young enough to benefit from the reform, discontinuously acquire more schooling. Husbands of these “young enough” females are more educated, and the age difference with most of these educated men decreases. The females’ own schooling, however, increases by more than their husbands’, so that the schooling difference between spouses declines. A framework is then introduced to examine whether the observed changes in spousal characteristics were associated with a change in complementarities and marital sorting patterns. The estimation of a so-called affinity matrix reveals that the reform has led to trade-offs in the marriage market, but not reinforced sorting on education.

Keywords: Marriage market, Education, Supply, Marital sorting
1 Introduction

Researchers now commonly recognize that education not only yields a return in the labor market, but also one in the marriage market (ex. Goldin 2006; Chiappori et al. 2009, 2017). That is, education not only affects future wages, but also the probability of getting married, the characteristics of the future spouse, and the surplus generated within marriage. Meanwhile, education expansion programs have taken place in numerous developing countries.\footnote{For instance, there have been school construction and expansion programs in Indonesia (Duflo 2001), Zambia (Ashraf et al. forthcoming), Sierra Leone (Mocan and Cannonier 2012), and Nigeria (Osili and Long 2008).} While the effects of school construction programs on schooling and labor market outcomes are well-studied\footnote{For instance, see Duflo (2001, 2004), Nishimura et al. (2008), and Ashraf et al. (forthcoming).}, our understanding of the impact of an aggregate shock to human capital on the marriage market is much more limited. Matching theories suggest that marriage outcomes depend both on preferences over a partner’s attributes and on the availability of men and women with these characteristics. Hence, education expansion programs can impact an individual’s marital outcome by affecting own educational attainment and by shifting the education distribution of others. Yet, few researchers have examined how matching patterns adjust when individuals’ attributes such as education are shifted.

This paper addresses this gap by asking the following research question: What is the impact of an aggregate shock to human capital on the marriage market? Stated differently, how does such a shock affect marriage probabilities and characteristics of spouses? And how do matching patterns adjust to such a shock over time? This research question will be addressed within the context of Zimbabwe, a country in southern Africa. Zimbabwe is a setting well-suited for this purpose as it allows us to overcome two commonly faced empirical challenges. First, a large, nationwide education reform was implemented in 1980 which provides us with an aggregate shock to human capital. The reform eliminated restrictions for advancing to secondary school, and thereby led to a dramatic increase in secondary school enrollment. The transition rate from the last grade of primary to the first grade of secondary education jumped from 27% in 1979 to 86% in 1980, making the progression from primary to secondary school the relevant margin in this context. Second, marriages prior to the reform are characterized by a large age difference between spouses. On average, wives were roughly 8 years younger than their husbands. In a given birth cohort, the reform could affect the education of both males and females. However, since females enter the
marriage market at an earlier age, the 1980 reform affects the education of future *wives* first. In particular, of the first female cohort exposed to the reform, only 5% married men young enough to also have benefited from the reform. Thus, this reform *initially* affected education levels of only one side of the marriage market.

The first part of the analysis exploits the age-specific exposure to the 1980 education reform in a regression discontinuity (RD) design to estimate the causal effect of a change in female education on marriage outcomes. The intuition underlying this RD strategy is that individuals who were still young enough for secondary school at the time of the reform discontinuously acquire more schooling for plausibly exogenous reasons. This quasi-random assignment then allows us to compare the marriage market outcomes of individuals “just young enough” to the outcomes of individuals “just too old” and to attribute any discontinuity in these outcomes to schooling. Using data from Demographic and Health Surveys, I find that females who were still young enough for secondary school when the reform was implemented discontinuously acquire more schooling. “Just young enough” females obtain 1.6 additional years of schooling relative to females “just too old” for the reform, where this increase is primarily driven by the progression from primary to secondary school.

In Zimbabwe, marriage is nearly universal and the 1980 education reform does not cause females to increasingly remain single. Conditional on entering into marriage, females who obtained more schooling as a result of the reform seem to desire an educated spouse. I find that husbands’ schooling is discontinuously higher when their wife was still young enough to benefit from the reform. Nonparametric two-stage least squares (2SLS) estimates indicate that wives who were still young enough for secondary school in 1980 acquire 1.15 more years of schooling and then marry husbands with 0.88 more years of schooling. The wife’s schooling, however, increases by more than their husband’s, leading to a decline in the schooling difference between spouses. On the one hand, this could indicate a preference for a partner with similar levels of schooling and stronger sorting on education; on the other hand, this could signal a shortage of educated men available for marriage. As the 1980 reform affected only one side of the marriage market, demand for educated husbands could exceed the available supply of educated men. A change in the age composition of married men partly explains the rise in husbands’ schooling - some educated husbands come

---

3In practice, it turns out that the 1980 reform only clearly shifted the education distribution of females.
from younger cohorts. Again, females could match with younger, educated males because they generate higher utility together or because not enough older, educated males were available. That is, both preference and supply forces could drive the observed changes in matching patterns and the regression framework does not allow us to properly distinguish between the two.

The second part of the analysis introduces a framework based on Dupuy and Galichon (2014) to address the questions left open by the regression discontinuity design. In this one-to-one matching model with transferable utility, men and women are characterized by multiple continuous attributes. This framework lends itself to the estimation of a so-called affinity matrix. This affinity matrix captures the strength of complementarities between attributes of spouses, while controlling for the marginal distributions of these characteristics. Each spouse is characterized by two attributes, age and education, resulting in a $2 \times 2$ affinity matrix that contains four cross-partial derivatives. Estimating this matrix for different cohorts of females allows us to examine whether the 1980 education reform not only affected spousal characteristics but also the degree of assortative mating and dimensions of sorting.

The estimates of the affinity matrix reveal that, while education is an important attribute in the marriage market in Zimbabwe, the 1980 reform did not reinforce assortative mating on education. Thus, the observed decline in the spousal schooling difference has not been associated with an increase in complementarity between education of spouses but is rather explained by supply constraints. That is, a shortage of educated men initially forced some females to be the more educated spouse in the couple. The 1980 reform appears to, however, have increased the substitutability between wife’s education and husband’s age for the first female cohorts exposed to it. That is, educated women now generate higher joint utility with younger husbands and some therefore willingly choose an educated husband more similar in age.

This work helps to shed light on the impact of human capital on marriage market outcomes. School expansion policies have been implemented in many developing countries. Yet, how such programs affect matching patterns and related outcomes is still understudied and not well-understood. In the developing world, labor market opportunities - and hence outside options to marriage - are
often limited for females. As a result, marriage plays an important role for female welfare and studying how schooling affects marital outcomes helps us better understand an essential determinant of female welfare. Schooling programs can also have implications for aggregate economic outcomes policy makers care about, such as inequality. In Zimbabwe, both couples’ labor market and household production are increasing in the education of spouses. Accordingly, implementing a reform that raises educational attainments could improve childhood mortality rates and other children’s outcomes and therefore be desirable. This conclusion, however, might not hold if the schooling expansion reinforces assortative mating on education.\footnote{School expansion programs could reinforce assortative mating by making parents’ education a more important input in the production of children’s human capital, as suggested by Chiappori et al. (2017). In Zimbabwe, there is evidence that the interaction between spouses’ education already mattered beyond the contribution of their individual educational attainments for the couples’ labor market production prior to the reform (see Figure A.1 and Table A.1).} If the reform increases the proportion of couples in which either both spouses are educated or uneducated at the expense of couples in which only one spouse is educated, aggregate inequalities could be exacerbated and have detrimental effects for the next generation. Examining the impact of school expansion programs on sorting thus adds to our understanding of possible (unintended) indirect impacts such programs might entail.

This paper makes contributions to several literatures. First, it adds to a recent but growing literature studying the impact of education on marriage market outcomes. Hener and Wilson (2018) exploit a compulsory schooling reform in the UK in a regression discontinuity (RD) design and find that treated women decrease the marital age gap to avoid marrying less qualified men. This is consistent with my finding that females exposed to the 1980 reform marry educated men who are more similar in age. However, by complementing RD based analysis with a theoretical framework plus estimation, my work allows us to assess implications for marital sorting more thoroughly. Studying a school construction program in Cameroon, André and Dupraz (2018) find that education increases the likelihood of being in a polygamous union for both men and women only because more educated men are more likely to be polygynous and spouses assortatively match on education. The most closely related work to this paper is by Zha (2019) who revisits an extensively-studied school construction program in Indonesia and uses a difference-in-difference design to study the effect of decreased secondary school attainment on marital outcomes. While we both find evidence for

\footnote{Only 40% of all Zimbabwean women worked for pay prior to the 1980 education reform, of which 30% worked in agriculture. Of the non-agricultural work, 55% was conducted at home and 56% was declared self-employment.}
complementarity between education of spouses, Zha (2019) suggests a complementarity between wife’s age and husband’s education, whereas this work indicates a complementarity between wife’s education and husband’s age. Moreover, this paper also contributes to a literature based on the idea that education cannot only yield a return in the labor market but also one in the marriage market (ex. Goldin 2006; Chiappori et al. 2009, 2017). This literature posits that education can impact marriage outcomes by affecting marriage probabilities, spousal characteristics, and the size and distribution of the generated marital surplus. In addition, this paper complements a literature on the impact of marriage market conditions on individuals’ outcomes. While most of the existing work focuses on the sex ratio (ex. Abramitzky et al. 2011; Angrist 2002; Charles and Luoh 2010), this paper focuses on the shift in the education distribution of one side of the marriage market. Finally, this paper also contributes to a large literature focusing on the identification and estimation of preferences in the marriage and other matching markets. The modeling of marriage as a competitive market goes back to seminal work by Becker (1973), while empirical advances were initiated by Choo and Siow (2006) and pursued, among others, by Chiappori et al. (2010), Fox (2010, 2011), and Galichon and Salanié (2015). The theoretical framework in this paper is a matching model with transferable utility and separable idiosyncratic preferences based on work by Dupuy and Galichon (2014) who extend Choo and Siow (2006) to account for multiple, possibly continuous attributes of individuals. The main novelty of this paper is to estimate the evolution of complementarities in the marriage market over time and to trace out whether a large shock to education led to changes in assortativeness.

The remainder of this paper is organized as follows. Section 2 describes marriages in Zimbabwe prior to the 1980 education reform and provides details on the reform. Section 3 describes the data used. Section 4 outlines the empirical strategy, presents results, and discusses identification, robustness and limitations. Section 5 introduces a framework aimed at overcoming these limitations. Section 6 discusses its empirical implementation. Corresponding estimation results are presented in Section 7 and discussed in Section 8. Section 9 concludes.
2 Context

2.1 Marriages Prior to the Education Reform

Prior to the education reform, with 98% of all females entering into marriage, marriage was nearly universal in Zimbabwe.\[^{6}\] Moreover, these women married at a relatively young age and often entered marriages with large age differences between spouses. Females too old for the reform, on average, got married at the age of 19 to an 8-years older husband.\[^{7}\] This considerable age difference makes Zimbabwe an appealing and unique setting to study how a large shock to education impacts marital outcomes. Exposure to the education reform is age-specific and in a given cohort, the reform could affect the education of both males and females. However, since females enter the marriage market at an earlier age, future wives are exposed to the reform first (relative to future husbands). From a marriage market perspective, this reform can therefore be viewed as an initial education shock to only one side of the market and can affect relative supplies of educated spouses. Prior to the reform, educated men outnumbered educated women in the marriage market and husbands had, on average, 1.4 more years of schooling than their wives. While education levels of spouses have been rising over time, the spousal schooling difference has remained fairly stable.\[^{8}\] As the reform creates an initial excess supply of educated women, clearing the marriage market might require trade-offs. That is, females might need to accept having a relatively less educated husband or (since education levels are rising over time) need to resort to younger cohorts of men.

The two major ethnic groups in Zimbabwe, Shona and Ndebele, share similar marriage practices. In particular, both ethnic groups practice the bride price custom - a transfer from the groom’s to the bride’s family at the time of marriage - and polygyny, a marital union between one man and multiple women. The choice of a spouse is an individual decision, but the family is involved in the bride price negotiations and can influence the date of marriage. Recent work by Ashraf et

\[^{6}\text{Zimbabwe is a patriarchal and patrilineal society, where females status is tied to marital status and number of children. To avoid formal marriage, females need sufficient resources. Thus, there is strong pressure to get married. Divorce is generally rare and highly stigmatized.}\]

\[^{7}\text{Sample of women aged 30-49. At the time, the Marriage Act specified a minimum age of marriage of 16 for females and of 18 for males (minimum age of 18 for both genders only as of January 2016). Figure A.4 shows that Zimbabwe is no outlier in terms of its spousal age difference in Southern and Eastern Africa. Zimbabwe and other countries in the region also share a similar distribution shape: the age difference distribution is skewed to the right, has a concentration around the mode, and a low representation of negative values.}\]

\[^{8}\text{Prior to the 1980 reform, both spouses completed less than primary school, on average. Figure A.2 in the appendix provides a graphical illustration.}\]
al. (forthcoming) found that female schooling in bride price ethnic groups is higher and more responsive to a reduction in the cost of schooling, as the bride price constitutes an additional (monetary) incentive for parents to invest in the education of their daughters. However, due to unavailable bride price data and lack of variation in its practice in Zimbabwe, no further account of this custom will be taken in this work. Fenske (2015) provides evidence that schooling reforms did not significantly alter the decline in the prevalence of polygyny over time\footnote{In Zimbabwe, only 11\% of married women born in 1980 are in a polygynous union as opposed to 18\% of wives born in 1965 (sample of married women 25-49).}, making polygyny a non-primary outcome of interest in this work.

Prior to the education reform, less than 5\% of females married while they were still in school, suggesting that schooling and marriage are not concurrent activities. Moreover, females obtain close to zero years of additional schooling once they are married. However, there is no strong evidence that females drop out of school in order to get married. The age at marriage is between 17 and 18 for females with zero years of schooling and increases with educational attainment. That is, females with primary education tended to marry between 18 and 19, while females who had started secondary school married after they turned 20. Since students in Zimbabwe complete primary school at the age of 14 and the most advanced level of secondary school at the age of 20, females appear not to be choosing between schooling and marriage. This conjecture is further confirmed by the observation that only a small fraction of females (less than 5\%) end school and start marriage at the same age. On average, for females with at least one year of schooling, 5 to 6 years elapsed between the end of their schooling and the start of their marriage\footnote{This assumes an school entry age of 7 and uninterrupted schooling.}. For females with secondary education, this time interval is shorter but still amounts to 3 to 4 years and only turns slightly negative for females with higher education.

\subsection{2.2 The 1980 Education Reform}

In April 1980, the newly elected government of Zimbabwe implemented a large nationwide education reform. The new government sought to break with the apartheid-like education system that prevailed during colonial times in Zimbabwe\footnote{See Atkinson (1972) and O’Callaghan and Austin (1977) for a history of Rhodesia’s education system and the policies dictating the quantity and quality of schooling Africans received. More than 95\% of the population are Black according to historical Census and DHS data.}. The colonial system had created severe school pro-

\footnote{In Zimbabwe, only 11\% of married women born in 1980 are in a polygynous union as opposed to 18\% of wives born in 1965 (sample of married women 25-49).}

\footnote{This assumes an school entry age of 7 and uninterrupted schooling.}

\footnote{See Atkinson (1972) and O’Callaghan and Austin (1977) for a history of Rhodesia’s education system and the policies dictating the quantity and quality of schooling Africans received. More than 95\% of the population are Black according to historical Census and DHS data.}
gression bottlenecks for Black students by deliberately underinvesting in Black education, limiting school places, charging high school fees and requiring a series of highly selective examinations.\textsuperscript{12} For instance, in 1976, for every 1,000 black school-age children, only 750 ever started school, of which only 60 transitioned into secondary education (Nhundu 1992).

The education system in Zimbabwe consists of primary education, secondary education and higher education. The official school entry age at the time was seven years.\textsuperscript{13} Children complete seven years of primary education which is followed by a total of six years of secondary education divided in three two-year levels: junior, middle, and advanced. Students typically enter junior secondary education at the age of 14. In order to progress from junior to middle secondary education, students are required to pass the O-level examinations. After an additional two years of advanced secondary education, students then take the A-level examinations and can proceed to higher education.

The 1980 education reform consisted of four main initiatives: (i) free and compulsory primary education, (ii) automatic grade progression, (iii) removal of enrollment age restrictions, and (iv) community support for education. While the reform led to an increase in school enrollment at all education levels, the greatest expansion took place in secondary education. Enrollment in all schools rose by 205\%, from approximately 886,000 to 2,699,000, between 1979 and 1985. Yet, over the same period, secondary school enrollment increased by 628\%, from 66,000 to 482,000 (Nhundu 1992) (see Figure A.5). Figure A.6 shows that the transition rate from the last grade of primary to the first grade of secondary education jumped from 27\% in 1979 to 86\% in 1980. Although primary school completion was far from universal prior to the reform, lifting restrictions limiting the advancement toward secondary school and instituting automatic progression from primary to secondary grades had much larger enrollment effects than introducing compulsory primary education.\textsuperscript{14}

\textsuperscript{12} The government spent 12 times more per primary school student in the “European” system than in the “African” system and nearly three times more at the secondary school level (Dorsey 1989). The construction of secondary schools for Black students was heavily restricted, so that only one new public secondary school was built between 1961 and 1972 (Zvobgo 1981). Moreover, whereas White students automatically proceeded to secondary school, Black students had to pass an exam at the end of primary education in order to progress to secondary education (Dorsey 1989).

\textsuperscript{13} The school entry age was lowered to six years in 1989.

\textsuperscript{14} The probability of completing at least primary education does not discontinuously increase for students of primary school age at the time of the reform, who would have potentially been forced to complete primary school. More specifically, there is no discontinuity in primary school completion rates for cohorts born between 1966 and 1972 who were in grades 1-7 at the time of the reform.
margin of this reform therefore is the progression from primary to secondary school.

This tremendous increase in demand for schooling was accompanied with an increase in schooling supply. The government built new schools and undertook reconstruction and expansion of existing facilities. Between 1979 and 1985, the overall number of schools increased by 111%. Figure A.7 shows that the number of primary schools grew by 76%, while with an increase by 586%, the growth in the number of secondary schools was much larger (Agüero and Bharadwaj 2014). Figure A.8 shows that the number of teachers rose as well and that the student-teacher ratio remained at reasonable levels. In order to accommodate the increased demand by an increase in schooling supply, a large share of the government’s budget was allocated to education. In the fiscal year 1980-81, 22.1% of the national budget was allocated to education. This was a considerable increase from the share of 11.6% in the fiscal year 1979-80. The budget share allocated to education remained high at about 17% until 1986-87, demonstrating that education clearly was a government priority at the time (Dorsey 1989). Moreover, this education reform was a nationwide reform that sought to provide a secondary school within walking distance of all rural children.\footnote{Given the large government budget allocated to education, schools were built at a high speed all over the country. Agüero and Ramachandran (2014) confirm that while some districts indeed opened schools sooner than others, these disparities were eliminated by 1983 (only three years after the reform). This results in a lack of spatial variation precluding us from pursuing a similar empirical strategy as Duflo (2001).}

Nonetheless, concerns have been raised that this rapid expansion led to a decline in educational quality. Especially in rural areas, where increases in school enrollment occurred faster than classrooms could be built or teachers be trained (Nhundu 1992).

3 Data

The data primarily used in this paper comes from the 1988, 1994, 1999, 2005, 2010 and 2015 survey waves of the Demographic and Health Surveys (DHS).\footnote{In Zimbabwe, DHS surveys are conducted by the National Statistics Agency ZIMSTAT.} DHS surveys are standardized, nationally representative surveys conducted in developing countries around the world. The main objective of these surveys is to provide policy-makers with information on fertility, family planning, maternal and child health, as well as on socioeconomic and marital status. Women are the main unit of analysis in these surveys, so that most of the analysis will be conducted from the perspective
of females. In each survey wave, a nationally representative sample of women aged 15-49 is interviewed. Pooling different survey waves results in a repeated cross-section. From 1994 onward, males aged 15-54 living in households with female respondents are interviewed as well. However, since men are only interviewed when they live in a household eligible for a female survey, the DHS provides us with a representative sample for women but not men. In order to have a nationally representative sample for females, all of the analysis considers the provided survey weights.

The main variables of interest are demographics, marital status and spousal characteristics. Demographics such as year of birth and years of schooling allow us to determine the exposure to and the effectiveness of the 1980 education reform, whereas marital status and spousal characteristics constitute our outcome variables. Summary statistics are provided in table A.2 in the appendix.

The sample is restricted to individuals born between 1950 and 1980. Since the first cohort affected by the education reform was born in 1965, this provides us with a 15-cohort window on each side of the age cutoff. Moreover, since we are interested in the effect of schooling on marriage market outcomes, we restrict our sample to individuals who likely have completed their education. More than 96% of females completed their schooling by the age of 20, so we restrict our sample to individuals aged 20-49 at the time of the survey who report that they are not currently still in school. After some minor data cleaning, this leaves us with a sample of 22,192 females. In order to reduce survivor bias, we will examine the first marriage outcomes of ever married women. In addition to currently married women, this also includes separated, divorced and widowed women. Out of 22,192 females, 19,831 (89%) have ever been married. While data on husbands’ education is inquired in all survey waves, data on husbands’ age only becomes available from the 1999 survey wave onward. Data quality appears to be reasonable. For instance, there is not much evidence of age heaping in husband’s age. 23.1% of observations of husbands’ age end in digits “5” or “0” as opposed to 20% by chance.

---

17Since the DHS surveys do not provide us with a random sample for males, we turn to the 2012 Census as a secondary data source for robustness. A 5% random sample is publicly available from IPUMS International.
18In 1994, all males aged 15-54 in 40% of households selected for female surveys were interviewed. In 1999, this proportion rose to 50%, while from 2005 onward males in all households (100%) were interviewed.
19Individuals with missing years of schooling (20 observations) or missing marital status (2 obs.) are dropped, and so are females who got married at the age of 10 or younger (129 obs.) or after the age of 40 (25 obs.).
4 Empirical Analysis

The age-specific exposure to the 1980 education reform is exploited to estimate the impact of education on marriage market outcomes. Individuals who were still young enough for secondary school at the time of the reform discontinuously acquire more schooling for plausibly exogenous reasons. Our causal estimates of interest then can be obtained by comparing the marriage market outcomes of individuals “just young enough” to the outcomes of individuals “just too old” for the reform.

4.1 Empirical Strategy

Our empirical strategy aims at capturing the following relationship between education and marriage outcomes:

\[ Y_i = \alpha + \rho S_i + \eta_i \]  

where \( Y_i \) denotes an individual \( i \)'s marital outcome and \( S_i \) represents individual \( i \)'s educational attainment. Thus, \( \rho \) is the parameter of interest. In order to obtain valid causal estimates of this parameter, we will exploit the 1980 education reform as a source of exogenous variation in schooling in a regression discontinuity (RD) design.\(^{20}\)

The relevant margin of this education reform consisted in progressing from primary to secondary school. In Zimbabwe, students typically enter secondary school at the age of 14. This results in an age-specific exposure to the reform based on whether children were still young enough for secondary school when the reform was implemented. For these “young enough” individuals, the probability of schooling attainment discontinuously increases for plausibly exogenous reasons. Given that potentially not all “young enough” children obtain more schooling as a result of the reform, we will use the discontinuity in the probability of more schooling as an instrument. This form of the RD design is called fuzzy RD and can be estimated via two-stage least squares (2SLS) regressions. In the first stage, we regress schooling attainment on an indicator for still being young enough for secondary school at the time of the reform. The resulting predicted values are then used as an explanatory variable for marriage market outcomes in a second stage regression. Marriage market outcomes via ordinary least squares (OLS) would yield biased estimates if unobserved individual characteristics such as ability or family attitudes affect both the education attainment and marriage decision.\(^{20}\)
outcomes of interest are: (i) marriage probabilities, (ii) education of spouse, (iii) spousal education difference, and (iv) spousal age difference. Data limitations prevent us from considering other potentially relevant outcomes such as measures of physical or financial attractiveness.

The cutoff birth cohort for “being young enough” is taken to be 1965. Individuals born in 1965 were 15 years old when the reform was implemented in 1980. While the “regular” secondary school cohort consists of individuals aged 14 in 1980 (born in 1966), individuals aged 15 in 1980 were also induced to acquire significantly more schooling. In the analysis that follows, the 1965 birth cohort will be included in the graphs, but excluded from the regression estimations. In Zimbabwe, the school year and the calendar year are aligned.

The baseline estimation will be implemented nonparametrically. The nonparametric estimation amounts to conducting an IV estimation in a small neighborhood around the discontinuity. Nonparametric estimation has the advantage that a nonlinearity is unlikely mistaken for a discontinuity and that the exact functional form of \( E[Y_{0i} \mid \text{Age at Reform}_i] \) matters little, where \( Y_{0i} \) is the potential outcome of individual \( i \) in the absence of the reform. Moreover, the validity of this nonparametric estimation strategy does not turn on the constant effects assumption. The size of the neighborhood will be chosen in a mean-squared-error optimal way and we will allow for different bandwidths on each side of the cutoff. Observations further away from the cutoff receive less weight by the use of a triangular kernel and robust standard errors are computed. The first stage takes the following form:

\[
E[S_i \mid \{\text{Born in 1965}\} < \text{Age at Reform}_i < \{\text{Born in 1965}\} + \delta] - E[S_i \mid \{\text{Born in 1965}\} - \delta < \text{Age at Reform}_i < \{\text{Born in 1965}\}] \simeq \pi
\]

where \( S_i \) denotes individual \( i \)'s years of schooling, \( \{\text{Born in 1965}\} \) is the cutoff birth cohort and point of discontinuity, and small \( \delta > 0 \) is the size of the neighborhood considered around this cutoff. In words, equation 2 compares the schooling of individuals just to the right of the birth year cutoff (“just young enough for secondary school”) to the schooling of individuals just to the left of the

\[21\] This could be due to the lifting of overage enrollment restrictions as part of the reform. Relative to the 1964 birth cohort, females born in 1965 obtain 0.9 more years of schooling while females born in 1966 acquire 1.6 additional years of schooling.

\[22\] Using data closer to the cutoff reduces bias but also precision.
cutoff (“just too old for secondary school”).

The reduced form does a similar comparison for marriage market outcomes $Y_i$:

$$E[Y_i \mid \{\text{Born in 1965}\} < \text{Age at Reform}_i < \{\text{Born in 1965}\} + \delta] - E[Y_i \mid \{\text{Born in 1965}\} - \delta < \text{Age at Reform}_i < \{\text{Born in 1965}\}] \approx \rho \pi$$

Again in words, equation (3) compares marriage outcomes of individuals “just young enough for secondary school” to those of individuals “just too old for secondary school”.

Therefore,

$$\lim_{\delta \to 0} \frac{\text{Reduced Form}}{\text{First Stage}} = \rho$$

The sample analog of (4) is simply a Wald estimator and captures a local average treatment effect (LATE). In particular, this estimator captures the causal effect of schooling on marriage market outcomes for the compliers - the individuals who acquire more schooling as a result of the reform. More specifically, compliers in this context are individuals who wanted to progress to secondary school, but were prevented from doing so in the pre-reform system. When given the chance to advance to secondary school, 86% of eligible individuals are induced by the reform to change their behavior and acquire more schooling. This suggests that the set of compliers is large and that the local average treatment effect (LATE) will be close to the average treatment effect (ATE).

In addition, as usual in IV estimation, we assume that, in a small neighborhood around the cutoff, an individual’s birth year affects his/her marriage outcomes solely through schooling attainment (IV exclusion restriction). However, while the exclusion restriction commits us to a specific causal channel, the assumed channel need not be the only one that matters in practice. Finally, this estimation strategy only yields causal estimates close to the cutoff, which in this context can be interpreted as the impacts for the cohorts first affected by the education reform.

4.2 Identification

The key identifying assumption underlying a regression discontinuity design is that there is nothing special about the birth cutoff, other than determining schooling. That is, close to the cutoff, whether an individual is born to the left or to the right of the age cutoff should be as good as random. Formally, we require $E[Y_{0i} \mid \text{Age at Reform}_i]$ and $E[Y_{1i} \mid \text{Age at Reform}_i]$ to be smooth
in $Age at Reform_i$ at the 1965 cutoff, where $Y_{0i}$ and $Y_{1i}$ denote the potential outcomes when not exposed and exposed to the reform, respectively. Given this quasi-random assignment, any discontinuity in marriage outcomes can then be attributed to schooling and average outcomes of individuals born just to the left of the cutoff can be used as a valid counterfactual for individuals born just to the right of it.

This identifying assumption, as usual, cannot be tested directly. However, two validity checks can provide evidence in support of this assumption. First, we can examine the density of $Age at Reform_i$ for bunching near the 1965 birth cohort cutoff. Such bunching would be a sign of manipulation of $Age at Reform_i$, the running variable, and suggest that $E[Y_{0i} | Age at Reform_i]$ is not smooth near the cutoff. The left side of Figure 1 depicts a histogram by year of birth, the equivalent to $Age at Reform_i$, and shows no sign of manipulation. This evidence is confirmed when testing for manipulation more formally using local polynomial density estimation techniques as proposed by Cattaneo et al. (2018). The result of this test is presented graphically on the right side of Figure 1. The null hypothesis of no systematic manipulation of the running variable around the cutoff cannot be rejected ($p$-value = 0.837). This is taken as evidence that exposure to the reform was as good as random. Moreover, given that nothing is tied to birth year or current age in the DHS survey, no strategic considerations or incentives to misreport should be at play.

A second validity check examines the behavior of pre-reform characteristics around the cutoff. Since pre-reform characteristics are unaffected by the reform, there should be no discontinuity in the conditional expectation function of these variables at the birth year cutoff. In this context, an individual’s height (in cm) and race (an indicator for being black) are arguably two characteristics that should not have been affected by an education reform. Figure 2 provides evidence that these two characteristics are indeed continuous around the cutoff of 1965, which a regression analysis also confirms more formally.

Beyond that, we can test for discontinuities at other “placebo” cutoff values. For instance,
Figure 1: Density of Age at Reform

Source: Zimbabwe DHS 1999. Sample of women 20-49. Left: Histogram by year of birth. Right: Local polynomial density estimator with 95% confidence intervals. Vertical line drawn at cutoff birth cohort of 1965. Females born before 1965 were too old to benefit from the 1980 education reform, while females born after 1965 were still young enough. No evidence of manipulation of the running variable, Age at Reform, near the cutoff.

we can examine whether educational attainment of individuals aged 10 or 20 years at the time of the reform discontinuously increases as well. In both cases, no evidence of discontinuities at these placebo values is found. Finally, Agüero and Bharadwaj (2014) confirm that there is no discontinuity at age 15 in other countries in sub-Saharan Africa where no major reform took place.

Hence, both validity checks and placebo tests provide strong evidence in support of the identifying assumption. It is important to note that 1980 is also the year in which Zimbabwe gained independence. This empirical strategy, however, relies on the fact that the education reform created a discontinuity that disproportionately affected individuals close to the birth year cutoff for secondary school. While plenty of other changes and reforms have occurred at the same time, our strategy is valid as long as these other reforms did not affect individuals differentially around our age exposure cutoff of interest. With the goal of reducing racial inequalities between the minority whites and majority blacks in all spheres of life, the new government also implemented land and agricultural reforms, health reforms, labor market reforms, tax and other reforms around 1980.

26The land reform aimed at redistributing land on a willing basis, which resulted in slow and limited redistribution.
However, none of these other reforms had a discontinuous impact on individuals aged 15 in 1980, who were still just young enough for secondary school.

The reforms implemented by the new government were met with high expectations and hopes for better lives. In response to demand from the electorate, the electoral manifestos of the three major parties had all agreed upon the need for free primary school education and major improvements in the health sector. The winning party’s manifesto had the goal of establishing free and compulsory primary and secondary education for all Zimbabwean children and abolishing racial and gender discrimination in the education system (Nhundu 1992). This was a reaction to high demand for greater provision of education by citizens who viewed education as key to jobs in the modern sector and therefore to upward mobility and a better standard of living (Dorsey 1989). The high

The agricultural reforms mainly expanded extension services to also include small-scale farmers in efforts to improve land. The health reforms aimed at improving access to health facilities by constructing and upgrading health institutions, recruiting and training additional health personnel, providing free health services and vaccinations. The labor market reforms meant to reduce wage disparities and discriminatory labor practices by introducing a minimum wage, requiring ministerial approval before firing a worker, and giving women the right to employment.

27 However, since the parties agreed on many policy issues, the election turned out to be rather decided by appealing to political credentials and ethnic identities of the parties’ leaders.
demand and expectations do not pose a threat to identification, but increase the set of compliers and thereby external validity of results. In this context, compliers are individuals who wanted to stay in school but previously could not due to the apartheid-like regime. The larger the set of compliers, the closer the reform’s estimated impacts will be to average treatment effects.

4.3 First Stage

*Females*

Females born after 1965 - and thus still young enough for secondary school when the education reform was implemented - discontinuously acquire more schooling. This is shown in Figure 3. In the left graph, each dot represents the average years of schooling for a given birth cohort to which 95% confidence intervals have been added. A vertical line has been drawn at the cutoff birth cohort of 1965. Females born to the left of the cutoff were 16 years or older when the reform was implemented in 1980 and thus too old for secondary school, while females born to the right of the cutoff were 14 years or younger and therefore still young enough to benefit.

Females born in 1966, just to the right of the cutoff, acquire 1.6 additional years of schooling relative to females born just to the left of the cutoff in 1964. As shown on the right in Figure 3, this increase in years of schooling is primarily driven by the progression from primary to secondary school. Females born in 1966 are 23 percentage points more likely to have started secondary school relative to females born in 1964. While the education of females at least weakly increases at all levels, the greatest jumps are recorded for starting and then completing up to four years of secondary school (see Figure A.9 and A.10 in the appendix).

This increase in education caused by Zimbabwe’s 1980 reform is considerable. In her seminal study of a school construction program in Indonesia, Duflo (2001) finds that an additional primary school per 1,000 children led to 0.12 to 0.19 more years of schooling for males; a 1.5% to 2.4% increase from a mean of 8 years of schooling. Zimbabwe’s education reform, on the other hand, increased education of “young enough” females by 1.6 years which represents a 27% increase from their slightly older counterparts’ average schooling of 6 years.

The slight upward trend in female education for cohorts born after 1960, as suggested by Figure 3 does not threaten the internal validity of our empirical strategy. For the subsequent 2SLS estimation to be internally valid, we require a strong instrument in the form of a clear discontinuity
Figure 3: First Stage of Females

![Graph showing the first stage of females' educational attainment.](image)

Source: Zimbabwe DHS 1988, 1994, 1999, 2005, 2010, 2015. Sample of women 20-49. Averages by year of birth. 95% confidence intervals have been added. Vertical line drawn at cutoff birth cohort of 1965. Females born before 1965 were too old to benefit from the 1980 education reform, while females born after 1965 were still young enough. Females born to the right of the cutoff discontinuously obtain more schooling.

In educational attainment around the cutoff birth cohort. This discontinuity is unequivocally present. The upward trend, however, reduces the set of potential compliers - individuals who wanted to progress to secondary school, but were prevented from doing so in the pre-reform regime. Since the reform’s impacts will be estimated for this set of individuals, this could raise concerns about the external validity of findings. Yet, given that the set of compliers is still much larger than in order settings, these concerns are not of the first order.

Males

In contrast to females, there is no first stage for males. As shown in Figure 4, while schooling of males appears to increase at a faster rate for cohorts born after 1960 and to level off for cohorts after

---

28 For instance, when compulsory schooling laws are exploited as natural experiments, the set of compliers is characterized by those who would drop out in the absence of the laws, but must remain in school under the new regime. Since these laws do not alter the behavior of those who already wanted to stay in school, Oreopoulos (2006) argues that most compulsory laws typically affect fewer than 10% of the population exposed to them.
born after 1965, no clear discontinuity can be detected.

Figure 4: First Stage of Males

However, given that average male schooling already exceeded primary education (7 years) for cohorts born after 1955, and given that the transition from primary to secondary school constituted the reform’s relevant margin, a missing first stage for males does not come as a real surprise. This is robust to using the 2012 Census (see Figure A.12).

Several potential explanations can be offered to rationalize the accelerated rate at which males born between 1960 and 1964 acquired schooling. First, this could be an indication of male overage enrollment due to a high value ascribed to education. That is, when the 1980 education reform lifted restrictions from secondary education, males who were technically too old for secondary school enrolled nonetheless. As Nhundu (1992) noted, colonial policies and practices to restrict Black educational advancement paradoxically created an even greater demand for education by Blacks as

---

29 In contrast, females born in 1964 or before completed less than primary school, on average. For males, the modal years of schooling switched from 7 years to 11 years for cohorts born in 1963 and after.

30 One initiative of the 1980 education reform consisted in removing enrollment age restrictions.
education was seen to be the only route to salaried employment and upward mobility. Presumably, this expected labor market return to education was higher for males who historically dominated waged employment in Zimbabwe than for females who were mostly confined to unwaged family work (Parpart 1995). Second, given the limited number of places in secondary school and the presumed higher expected return to secondary education, educational priority might have been given to males prior to 1980 and thereby lowered males’ costs of schooling relative to females’. Third, the national liberation war leading up to independence might have disproportionately disrupted the education of males born between 1960 and 1964. These males might then have resumed their schooling once the 1980 reform had lifted restrictions from secondary education. Finally, the observed change in the rate of educational attainment could be due to measurement error in either age or years of schooling.

Hence, while the 1980 reform shifted the education distribution of entire cohorts of females, it does not constitute a valid natural experiment to study the impact of male schooling on marriage outcomes. The remainder of this analysis will thus be conducted from the perspective of females.

4.4 Reduced Form and 2SLS

In this section, we study how an increase in female education affected their marriage probabilities and conditional on being married, characteristics of their husbands such as age and education.

Marriage Probabilities

Marriage in Zimbabwe is nearly universal and the 1980 education reform did not cause females to suddenly remain single. As shown in Figure 5, while marriage probabilities decline over time and more females tend to never get married, this does not happen discontinuously so around the cutoff birth year of 1965. These average marriage probabilities also do not mask any heterogeneous

---

31 According to Papart (1995), although women increasingly entered the waged labor force following independence, men continued to dominate waged employment and disproportionately gained from access to employment. In 1981, African women in Zimbabwe held 1.9% of waged jobs in production, 2.5% in agriculture/animal husbandry/forestry, 7.2% in services, 8.5% in clerical employment, and 16.6% of waged professional and technical jobs.

32 In colonial times, 50% of Black students reaching the last grade of primary education were precluded from proceeding to secondary education (Nhundu 1992).

33 The war forced thousands of students to flee to neighboring countries, where they were settled in refugee camps and received only a limited provision of education. The war also destroyed a total of 1,600 schools, leaving 483,000 children without education. The schools that still operated had far more children than they could accommodate. Moreover, the disruption of the local economy made it difficult for many parents to pay the school fees (Zvobgo 1981).
This lack of discontinuity in marriage rates motivates focusing on changes in spousal characteristics hereafter. That is, do future wives who were still young enough to benefit from the 1980 reform and acquired more education as a result, marry husbands with different characteristics? However, while overall marriage rates are not affected, the reform does cause females to marry at a slightly older age.\textsuperscript{35}

\textit{Husband’s Years of Schooling}

Females who were still young enough to benefit from the reform discontinuously marry more educated husbands. Figure 6 shows a clear discontinuity in husbands’ years of schooling when their \textit{wife} is born after 1965. A dot represents husbands’ average years of schooling for a given birth cohort of ever married women.

\textsuperscript{34}Educational attainments of never married females are generally higher than those of ever married females.

\textsuperscript{35}According to the nonparametric 2SLS estimation, females who acquire 1.2 more years of schooling (statistically significant at the 1\% level) marry at a 0.5 years older age (statistically significant at the 10\% level). See \textsection A.13 for additional graphical reduced form evidence.
Females born prior to 1965 were married to husbands who had, on average, completed just over primary education, whereas females born after 1965 are married to husbands who completed at least one year of secondary education. Husbands of females born after 1965 are also discontinuously more likely to have started secondary school. The nonparametric 2SLS estimation confirms the visual reduced form evidence presented here. Married females who were still young enough for secondary school in 1980 acquire 1.15 more years of schooling (first stage), statistically significant at the 1% level. Relative to women too old to benefit from the reform, these women then marry men with 0.88 more years of schooling (2SLS estimate), again statistically significant at the 1% level. Husbands of these women are also 11 percentage points more likely to have started secondary school (2SLS estimate). Hence, women who acquired more schooling as a result of the education reform went on to marry more educated men. The probability of marrying an educated man had already been increasing in women’s education prior to the reform. For instance, the probability of marrying a
man with some secondary education was three times higher for women who had proceeded on to secondary school relative to women who had completed at most primary school.\textsuperscript{36} Therefore, the reform could have led to an increase in husband’s years of schooling simply by raising the proportion of females with some secondary education.

However, when we plot the educational attainment of husbands and wives side by side, as done in Figure 7, we notice that the reform caused married women’s education to increase by more than married men’s. As a result, the average schooling difference between spouses, defined as the average over husband’s years of schooling minus wife’s years of schooling, discontinuously declines at the cutoff. Females who acquire 1.38 more years of schooling marry husbands with only 1 more year of schooling, i.e. the spousal schooling difference decreases by 0.38 years (2SLS), statistically significant at the 1% level. Both preference and supply forces could drive this decline. On the one hand, a preference for a partner with similar levels of schooling could lead to a lower schooling difference; on the other hand, an excess supply of educated women (or a shortage of educated men) could cause the schooling difference between spouses to decline. Some evidence suggests that supply forces dominate. First, while the schooling difference drops at the cutoff, it appears to rise and stabilize for female cohorts born after 1970, as shown on the right of Figure 7. In addition, when we compare married women’s to married men’s years of schooling, we observe an increase in the proportion of couples in which the wife is the more educated spouse that exactly offsets the decline of couples in which the husband is more educated. This, however, is only observed for the first few female cohorts affected by the reform. The proportion of couples with the wife as the more educated spouse already starts decreasing three cohorts after the cutoff (see Figure A.14). This fast reversion suggests that a shortage of educated candidate husbands was in part responsible for the observed decline in the spousal schooling difference and that this decline was an initial margin of adjustment to clear the marriage market.

Moreover, it is worth pointing out that, prior to the reform, some educated men married uneducated women, whereas the reverse was much less common.\textsuperscript{37} This is consistent with an excess

\textsuperscript{36}The probabilities amount to 76% and 21%; based on cohorts of wives born between 1950 and 1964 who were too old to benefit from the reform. As a result, married men’s years of schooling had been, on average, 4.3 years higher when their wife had some secondary education (10.1 years compared to 5.8 years).

\textsuperscript{37}For instance, 54% of men with some secondary education married a woman with primary education or less, while the corresponding share of educated women amounted to merely 25%. Consequently, the proportion of couples in which only the husband had some secondary education (17%) was much larger than the proportion of couples in which only the wife did (4%).
supply of educated men who receive their marginal contribution to a marriage with an uneducated woman - regardless of the education level of their actual wife - and are therefore indifferent between marrying an educated or uneducated woman. The reform, however, appears to have considerably reduced this relative excess supply of educated men in the marriage market.\textsuperscript{38} Directly after the reform, men with some secondary education are less likely to “marry down” and some men with less than secondary education are actually more likely to “marry up”.\textsuperscript{39}

Given the supply considerations just described, the question where these husbands with dis-

\textsuperscript{38} The proportion of couples in which only the wife has some secondary education initially more than doubled. However, it then quickly started to revert back to its pre-reform level (see Figure A.15). The proportion of couples in which only the husband went to secondary school mirrors this evolution.

\textsuperscript{39} Prior to the reform, 46% of men with some secondary education married a woman who also had some secondary schooling, whereas this proportion rises to 73% when the first five female cohorts exposed to the reform are considered. Only 7% of men with less than secondary education married a woman who had started secondary school prior to the reform, whereas this proportion increases to 30% immediately following the reform. As expected, the reform also led to a discontinuous and significant rise in the proportion of couples in which both spouses had some secondary education.
continuously more schooling come from arises all the more. The 1980 reform clearly shifted the
education distribution of females but not males, in a setting, where husbands were roughly eight
years older than their wives, on average. Yet, we observe a discontinuous rise in husband’s years of
schooling at the cutoff. One explanation, to which we turn next, could be that the age composition
of married men changed, i.e. that educated husbands come from younger cohorts. An alternative
explanation could be a change in the education composition of unmarried men. That is, uneducated
men could have been forced into singlehood. However, while uneducated men are more likely to
remain single over time (see Figure A.16), they account for an insufficiently small proportion of
men to quantitatively explain the increase in husband’s education observed in Figure 6.

40 Husband’s Age

Prior to the reform, women married an eight years older man on average. If women first
exposed to the reform continued to adhere to this age difference, their husbands would be born in
1958, where levels of male education were still low. At first glance, the observed discontinuity in
husbands’ schooling therefore seems puzzling. One explanation is that not only the education but
also the age composition of married men changed. That is, given the time trend in male schooling
achievements, it is plausible that educated husbands are relatively younger.

There is no evidence that the 1980 education reform caused females to discontinuously marry
younger men on average. However, while the average age difference between spouses appears to
be declining over time, there is no clear discontinuity around the 1965 birth year cutoff (Figure
A.17). The regression estimation confirms this visual evidence. Females who acquire 1.6 more years
of schooling marry men with an age difference lower by -0.29 years (2SLS), which is statistically
insignificant at all conventional levels.

Yet, there is some evidence that this zero average effect masks a shift in the age distribution of
married men. The first female cohorts exposed to the reform are more likely to marry men who
are at most five years older than them. Interestingly, these relatively younger spouses drive most
of the observed increase in husbands’ schooling. Figure 8 depicts the proportion of husbands with

40 The proportion of husbands with some secondary education was estimated to increase by 0.11 at the cutoff.
Uneducated men start to constitute less than 10 percent of all men in 1953 and their share continuously diminishes
from there onward. While several entire cohorts of such men would have to remain single, in practice their marriage
rates do not fall below 85% over the relevant time period.

41 See Figure A.18 in the appendix for some graphical evidence and Table A.3 for formal estimation results.
some secondary education who are at most five years older than their wives. These husbands account for nearly two thirds of all married men with some secondary education immediately after the reform but less than half prior to the reform. Although schooling levels of married men increased across the board, husbands more similar in age to their wife are responsible for most of the increase in married men’s educational attainment. Thus, there is some evidence that the more educated husbands are also younger.

Taken together, this implies a potential interaction between age and education in the marriage market. Again, both preference and supply forces could be at work. On the one hand, females could have a preference for an age difference based on their partner’s education level. For instance, females could prefer a young, educated man to an old, educated one; but an old, uneducated man to a young, uneducated one. On the other hand, females could only have a preference over their partner’s educational attainment and enter marriages with younger, educated men due to

---

42 Figure A.19 in the appendix provides a more fine-grained age decomposition.
supply constraints, i.e. due to a lack of older, educated men. Given the present evidence, both forces are plausible. The preference force is supported by the observation that, even prior to the reform, for given female education the age difference with her husband varies with his educational attainment. Specifically, the spousal age difference is decreasing in the husband’s schooling.\footnote{Prior to the reform, given the wife had some secondary education, the average age difference amounted to 5.6 years when her husband also had been to secondary school (14\% of couples) but to 8.1 years if he had not been (4\% of couples). When only the husband had started secondary education the age difference was 7.1 years (17\% of couples), whereas it was 8.8 years when none of the spouses had been to secondary school (65\% of couples).} That husbands’ schooling appears to lower the age difference by more than wives’ further suggests that educated husbands always tended to be younger. This also suggests that age by itself is not an important attribute for sorting in the marriage market in Zimbabwe. In contrast, the observation that the share of educated husbands similar in age to their wife declines as time passes (and more older, educated men become available) supports supply forces. In spite of rising levels of male education, females later exposed to the reform do not continue to marry young, educated men in the same proportions. For instance, while husbands who are at most five years older than their wives constituted nearly two thirds of all husbands with some secondary schooling directly after the reform, they account for 10 percentage points less a few years later. As a result, the average age difference between spouses does not fall below six years - even when, after 15 years of exposure to the reform, nearly 80\% of husbands have some secondary education.

In sum, the evidence presented in this section suggests that the 1980 education reform increased the relative supply of educated women in the marriage market and that this led two at least two observed changes in matching patterns. Some of the females initially affected by the reform responded by being the more educated spouse in the marriage, whereas others chose an educated, but younger husband. Both preference and supply forces could be behind these adjustments.

4.5 Robustness

All results presented thus far are robust to alternative nonparametric and parametric specifications. Evidence that the main results are robust to alternative bandwidth selection methods, exclusion of survey weights, controlling for urban/rural status, alternative variance estimation, and clustering of standard errors is provided in Tables A.5-A.7 in the appendix. Evidence of robustness to parametric estimation using linear, quadratic and cubic polynomials and different windows around the birth
year cutoff (15 years, 10 years, 5 years) is given in Tables A.8-A.10. Since the 1980 reform primarily promoted access to secondary school for Black Zimbabweans, robustness to restricting the sample to ever married Black women is provided in Table A.11. However, with a population share of over 98%, Black Zimbabweans constitute the vast majority of the population. Moreover, results are robust to restricting the sample to women aged 25-49 (Table A.12) or to currently married monogamous women (Table A.13).

4.6 Discussion and Limitations

By exploiting a unique setting, this paper provides valuable and novel insights into how a one-sided shock to spousal attributes affects marriage outcomes. The 1980 education reform in Zimbabwe shifted the education distribution of entire cohorts of females but not males, making it possible to isolate the impact an increase in female education had on matching patterns. In addition, the large spousal age difference which prevailed prior to the reform allows us to examine how this shock propagated over time, i.e. how the marriage market adjusted after the attributes of one side were shifted. School construction programs and school reforms are widespread in developing countries. Yet, while their impact on schooling outcomes is well-studied, our understanding of their impact on marriage outcomes and matching patterns is still very limited.

The empirical analysis presented so far found that the 1980 education reform caused females who were still young enough for secondary school to acquire more schooling and that these females then went on to marry more educated husbands. In addition, evidence has been provided that these more educated husbands are also relatively younger. This could be an indication of married women trading off age and education of their husbands, i.e. that they are willing to choose a younger spouse when this spouse is also more educated. One should note that while an educated husband generally offers the benefits of higher labor market and household production, marrying a young husband might not always be desirable. While a smaller age difference could imply a greater degree of equality within the couple, a younger husband - who potentially has lower earnings and fewer assets accumulated - could also lead to lower material well-being. Thus, the effect of a husband’s

44 Of the first female cohort fully exposed to the reform, only 5% married men young enough to also have benefited from the reform according to the age exposure cutoff.

45 Other work has focused on the impact of schooling expansion policies on standard monetary outcomes, such as wage, income and wealth, and on health outcomes.

46 Ex. higher child quality.
age on his wife's welfare is not unambiguous.

However, this empirical approach also comes with several limitations. The regression analysis focuses only on one characteristic at a time and thereby fails to properly account for the possibly multidimensional nature of matching, i.e. that attributes of spouses are traded off. Furthermore, by virtue of the regression discontinuity design, we estimate effects close to the point of discontinuity. These estimates are most valid for the cohorts initially affected by the reform and therefore, can be viewed as estimates of the immediate effects of the reform. However, their validity is reduced as we move further away from the birth cohort cutoff to analyze transition dynamics. Beyond that, this approach remains silent on whether educated females now match with younger, educated males because they generate more surplus together or because there was a shortage of older, educated males. That is, the regression framework does not allow us to properly distinguish between preference and supply effects. Finally, conclusions about changes in assortativeness and dimensions of matching are also outside of this framework.47

In order to overcome these limitations, a framework that lends itself to a structural estimation will be introduced in the next section. This framework will take into account that matching and sorting occurs on multiple attributes and will control for the marginal distributions of attributes. Therefore, we will be able to examine whether the changes in spousal characteristics observed thus far were accompanied by a change in complementarities and dimensions of marital sorting.

5 Framework: Dupuy and Galichon (2014)

The framework based on Dupuy and Galichon (2014) lends itself to the estimation of a so-called affinity matrix. This affinity matrix captures the strength of complementarities between observable attributes of spouses and determines the dimensions on which sorting occurs. Estimating this affinity matrix for cohorts of females born before and after 1965, the birth year that determines exposure to the 1980 education reform, will allow us to assess whether and how this reform affected assortativeness and marital sorting over time.

47The sign and strength of complementarities cannot simply be inferred from the variance-covariance matrix of matches since a positive complementarity in one pair of attributes that dominates a negative complementarity in another pair of attributes could lead to positive correlation that is mistaken for positive complementarity in both pairs of attributes.
Based on Dupuy and Galichon (2014), this is a one-to-one, bipartite matching model with transferable utility. Men are characterized by a vector of observable attributes $x \in \mathcal{X} = \mathbb{R}^{d_x}$, while women are characterized by a vector of observable attributes $y \in \mathcal{Y} = \mathbb{R}^{d_y}$. These observable attributes could be characteristics such as age, education and height, but also personality traits. The probability distributions of attributes of matched men, $P$, and matched women, $Q$ are treated as exogenous. This has the implication that the estimation will be based solely on observed couples. Moreover, the distributions $P$ and $Q$ are assumed to be centered and to have densities $f$ and $g$, respectively. That is, Dupuy and Galichon (2014) represents an extension to multiple continuous attributes from multiple discrete attributes in Choo and Siow (2006).

A matching is the probability density $\pi(x, y) \in \mathcal{M}(P, Q)$ that a couple with attributes $(x, y)$ forms from the matched population.

$$\mathcal{M}(P, Q) = \left\{ \pi : \pi(x, y) \geq 0, \int_{\mathcal{Y}} \pi(x, y)dy = f(x), \int_{\mathcal{X}} \pi(x, y)dx = g(y) \right\}$$

When a man $m$ with attributes $x_m = x$ and a woman $w$ with attributes $y_w = y$ match, they generate joint match utility: $\Phi(x, y) + \varepsilon_m(y) + \eta_w(x)$. This joint utility consists of two additive components: (i) the deterministic match utility based on observable characteristics $\Phi(x, y)$, and (ii) the random match utility based on unobserved “sympathy shocks” $\varepsilon_m(y) + \eta_w(x)$. As is standard in this literature, the shocks $\varepsilon_m(y)$ and $\eta_w(x)$ drawn by individuals are assumed to be distributed according to i.i.d. centered extreme value type I distributions with scaling parameter $\sigma^2$. The deterministic surplus $\Phi(x, y)$ is endogenously shared between spouses and Choo and Siow (2006) have shown that, given these distributional assumptions on the sympathy shocks, this sharing is given by: $\Phi(x, y) = U(x, y) + V(x, y)$, where $U(x, y)$ is the systematic part of the joint utility obtained by a man $x$ who matched with a woman $y$. In order to give rise to a continuous logit framework, two additional assumptions are required. First, we need to assume that each individual only has access to a set of “acquaintances” from which potential partners will be considered, which is a random subset of the entire population of the opposite sex. Acquaintances of men and women are indexed by $k \in \mathbb{N}$ and $l \in \mathbb{N}$, respectively. Each acquaintance is characterized by observable

---

48 Since the baseline framework is identical, I only focus on elements key to understanding the mechanism and estimation procedure, and refer the interested reader to the original paper.

49 Dupuy and Galichon (2014) show that incorporating singles leaves the analysis unchanged, which will be a consequence of the independence of irrelevant alternatives property of the logit model.

50 Thus, match utility at equilibrium only depends on observable attributes and individuals will be indifferent between partners with the same observable attributes.
attributes \((y_k^m, x_l^n)\) and associated with a random sympathy shock \((\varepsilon_k^m, \eta_l^n)\). Second, \(\{(y_k^m, \varepsilon_k^m), k \in \mathbb{N}\}\) and \(\{(x_l^n, \eta_l^n), l \in \mathbb{N}\}\) are assumed to be the points of a Poisson process on \(\mathcal{Y} \times \mathbb{R}\) with intensity \(dy \times \exp(-\varepsilon)d\varepsilon\) and of a Poisson process on \(\mathcal{X} \times \mathbb{R}\) of intensity \(dx \times \exp(-\eta)d\eta\), respectively. Loosely speaking, this second assumption is required to give rise to the conditional choice probabilities and independence property of a continuous multinomial logit model.

Each individual chooses a partner so as to maximize his or her own payoff, which in equilibrium maximizes total utility. To be precise, the equilibrium matching \(\pi\) maximizes the social gain:

\[
\max_{\pi(x,y) \geq 0} \int_{\mathcal{X} \times \mathcal{Y}} \Phi(x, y) \pi(x, y)dxdy - \sigma \int_{\mathcal{X} \times \mathcal{Y}} \log \pi(x, y) \pi(x, y)dxdy
\]

subject to

\[
\int_{\mathcal{Y}} \pi(x, y)dy = f(x) \quad [u(x)] \\
\int_{\mathcal{X}} \pi(x, y)dx = g(y) \quad [v(y)]
\]

where \(\sigma\) reflects the intensity of the trade-off between sorting on observed attributes, \(\int_{\mathcal{X} \times \mathcal{Y}} \Phi(x, y) \pi(x, y)dxdy\), and on unobserved attributes, \(\int_{\mathcal{X} \times \mathcal{Y}} \log \pi(x, y) \pi(x, y)dxdy\). A higher \(\sigma\) signifies more unobserved heterogeneity and pulls the solution closer to random matching. Maximizing this social gain is subject to two “accounting” constraints, which are associated with the Lagrange multipliers \(u(x)\) and \(v(y)\). Loosely speaking, if we fix the characteristics of men at \(x\) and sum over all women \(y \in \mathcal{Y}\) these men are matched with, we should obtain the density of men with these characteristics, and analogously for women. Expressing problem \([5]\) as a Lagrangean and taking first-order conditions w.r.t. \(\pi(x, y)\) yields the equilibrium matching (for any \(x \in \mathcal{X}\) and \(y \in \mathcal{Y}\)):

\[
\pi(x, y) = \exp \left( \frac{\Phi(x, y) - u(x) - v(y)}{\sigma} \right)
\]

The affinity matrix is introduced into this framework by assuming the deterministic match utility \(\Phi(x, y)\) takes the following quadratic form: \(\Phi^A(x, y) = x^T A y\), where \(A\) is the \(d_x \times d_y\) affinity

\[\text{Rules out any systematic sympathy shocks, i.e. sympathy shocks that are correlated across observable attributes.}\]
This affinity matrix is our object of interest as it captures the strength of mutual attractiveness between attributes. Specifically, because one has: \( A_{ij} = \partial^2 \Phi(x, y) / \partial x_i \partial y_j \). \( A_{ij} \) therefore captures how the (marginal) gain in joint utility from increasing man’s \( x_i \)th attribute evolves as woman’s \( y_j \)th attribute increases. When \( A_{ij} > 0 \), then attributes \( x_i \) and \( y_j \) are complements in generating joint utility, which will give rise to positive assortative matching along these attributes. That is, all else equal, men and women with high \( x_i \) and \( y_j \) match together, as do men and women with low \( x_i \) and \( y_j \). On the other hand, attributes \( x_i \) and \( y_j \) are substitutes in creating joint utility when \( A_{ij} < 0 \), giving rise to negative assortative matching. Hence, \( A \) captures the strength of complementarity/substitutability of attributes in joint utility creation.

6 Empirical Implementation

This section describes the procedure used to estimate the affinity matrix \( A \). Estimating this matrix for different birth cohorts of females allows us to examine whether the 1980 education reform not only affected spousal characteristics but also complementarities and dimensions of marital sorting.

Specifically, the affinity matrix \( A \) is estimated as follows: each value of \( A \) yields an equilibrium matching distribution \( \pi \). Given \( \pi \), the affinity matrix \( A \) can be obtained via a moment matching procedure, where the relevant moment will be the covariance between observed attributes of men and women. In practice, the implementation takes two steps. The first step consists in a fixed-point algorithm to compute \( \pi \), derived from the first-order conditions of the equilibrium match problem (5). In a second step, proximal gradient descent is used to find the low-rank \( A \) that implies a covariance that best matches the covariance between attributes observed in the data.

Substituting the equilibrium matching (6), which was obtained by taking first-order conditions of (5), into the two associated “accounting” constraints yields:

\[
\begin{align*}
 u(x) &= \sigma \log \left[ \frac{1}{f(x)} \int_y \exp \left( \frac{\Phi^A(x, y) - v(y)}{\sigma} \right) dy \right] \\
 v(y) &= \sigma \log \left[ \frac{1}{g(y)} \int_x \exp \left( \frac{\Phi^A(x, y) - u(x)}{\sigma} \right) dx \right]
\end{align*}
\]

When these two equations are combined, a fixed-point map is obtained on which we can iterate.

52 Recall \( d_x \) is the number of observed attributes of men, while \( d_y \) is the number of observed attributes of women. Dupuy and Galichon (2014) show that \( \Phi(x, y) \) could be allowed to be any polynomial of \( x \) and \( y \), making the quadratic parametrization a less restrictive assumption than it might seem. However, when dealing with continuous attributes, some parametrization is needed to avoid a missing data problem in the estimation.
The iterative projection fitting procedure (IPFP) represents an efficient algorithm for determining its numerical solution, which will be our estimate of \( \pi^A \). The procedure only requires initial guesses for \( A \) and \( v(y) \) and has been shown to have strong convergence properties.\(^{53}\)

Given \( \pi^A \), we then attempt to find the value of \( A \) such that the model implied covariance between observable attributes of spouses, \( \mathbb{E}_{\pi^A}[X_i Y_j] \), matches the covariance observed in the data, \( \mathbb{E}[X_i Y_j] \). Since the rank of the affinity matrix \( A \) will shed light on the number of dimensions on which sorting occurs, we would like to find a low-rank \( A \) that best approximates the data, i.e. we would like to subject the moment matching procedure to a rank constraint. In practice, this is implemented via nuclear norm regularization, resulting in the following optimization problem:

\[
\min_A \max_{\pi \in \mathcal{M}(P,Q)} \left\{ \mathbb{E}_\pi [x' A y] - \sigma \mathbb{E}_\pi [\log \pi(x, y)] - \mathbb{E}_{\hat{\pi}} [x' A y] \right\} - \mathbb{E}_{\tilde{\pi}} [x' A y] - \gamma \|A\|_*
\]

where \( \|A\|_* \) denotes the nuclear norm of \( A \) and \( \gamma \) is a penalty parameter. This objective function can be decomposed into two components. \( g(A) \), which is convex and differentiable, with \( \nabla g(A) = \mathbb{E}_{\pi^A}[X_i Y_j] - \mathbb{E}_{\hat{\pi}}[X_i Y_j] \) by the envelope theorem, and \( h(A) \), which is convex but not necessarily differentiable. Thus, we face a convex optimization problem for which existence and uniqueness are well understood under very general conditions on \( P \) and \( Q \).\(^{54}\) One method to compute a solution to this problem numerically is proximal gradient descent:

\[
A^{(k)} = prox_{t_k} \left( A^{(k-1)} - t_k \nabla g(A^{(k-1)}) \right)
\]

where applying the proximal operator \( prox(\cdot) \) consists in applying a soft-threshold operator to the singular values of the argument \( A - t \nabla g(A) \). A singular value decomposition of \( A - t \nabla g(A) \) yields \( U \Sigma V^T \), where \( \Sigma \) contains the singular values \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d \geq 0 \). The soft-threshold operator is then applied to these singular values contained in \( \Sigma \):

\[
S_\gamma(\lambda_i) = \begin{cases} 
\lambda_i - \gamma t_k & \text{if } \gamma < \lambda_i \\
0 & \text{if } -\gamma \leq \lambda_i \leq \gamma 
\end{cases}
\]

\(^{53}\)For more details on the computation, see Dupuy and Galichon (2014) or Rüschendorf (1995).

\(^{54}\)For more details and proof see Dupuy and Galichon (2014).
Substituting in for the proximal operator in (7) yields an expression for \( A \) on which we can iterate:

\[
A^{(k)} = U^{(k-1)}S_\gamma(\Sigma^{(k-1)})V^T^{(k-1)}
\]  

(9)

This estimation approach allows us to identify \( \frac{\partial^2}{\partial x \partial y} \Phi(x, y) \) ("mutual attractiveness"), which given our parametrization coincides with our object of interest: \( A = \frac{\partial^2}{\partial x \partial y} \Phi(x, y) \). However, we could identify neither \( \frac{\partial}{\partial x} \Phi(x, y) \) nor \( \frac{\partial}{\partial y} \Phi(x, y) \) ("absolute attractiveness"). That is, we can identify whether educated men and educated women attract each other, but not whether educated men and women are more desirable by themselves. This follows from only being able to identify the deterministic match utility \( \Phi^A(x, y) = x^T Ay \) up to a separatively additive function when we conduct our estimation on the matched population. In particular, observing \( \pi(x, y) \) leads to the identification of the deterministic match utility going to each spouse, \( U(x, y) \) and \( V(x, y) \), only up to an additive term, \( c(x) \) and \( d(y) \), respectively. Since \( \Phi(x, y) = U(x, y) + V(x, y) \), observing \( \pi(x, y) \) does not allow us to distinguish between \( \Phi(x, y) \) and \( \Phi(x, y) + c(x) + d(y) \) as they give rise to the same equilibrium matching.

Furthermore, this estimation procedure allows us to identify our object of interest, the affinity matrix \( A \), while achieving asymptotic statistical efficiency. It can be shown that:

\[
\sqrt{n} (\hat{A} - A) \xrightarrow{d} N(0, F^{-1})
\]

where \( n \) is the number of couples, \( \hat{A} \) is our estimator of \( A \) and \( F \) is the Fisher information matrix defined by:

\[
F_{ij} = E_\pi \left[ \frac{\partial \log \pi(X, Y)}{\partial A_{ij}} \frac{\partial \log \pi(X, Y)}{\partial A_{kl}} \right]
\]

In the empirical implementation, age and education of spouses are taken to be the observable attributes, resulting in a \( 2 \times 2 \) affinity matrix\(^{56}\). Each entry of this affinity matrix corresponds to a cross-partial derivative. Attributes have been standardized (mean 0, sd 1), so that the magnitude of coefficients is directly comparable across attributes. Diagonal entries can be interpreted as the change in joint utility when the given attribute of both spouses is increased by 1 standard deviation. The diagonal elements \( Age-Age \) and \( Education-Education \) therefore indicate how valuable

\(^{55}\)For more details and proof see Dupuy and Galichon (2014).

\(^{56}\)This choice is dictated by data limitations. However, while Dupuy and Galichon (2014) use a set of 10 attributes for each spouse, they find that education is the single most important attribute in the marriage market.
assortative mating on these attributes is for the creation of joint utility. The off-diagonal elements \textit{Wife Age-Husband Education} and \textit{Wife Education-Husband Age}, in turn, indicate whether and how spousal attributes are traded off. For instance, a positive entry in \textit{Wife Education-Husband Age} would indicate that increasing a married woman’s education increases the couple’s joint utility when their husbands are older. Whether males and females have similar preferences for attributes can be ascertained by comparing the sign and magnitude of the off-diagonal elements \textit{Wife Age-Husband Education} and \textit{Wife Education-Husband Age}.

The estimation will be conducted on a sample of matched couples and for different birth cohorts of \textit{wives}. Specifically, the affinity matrix will be separately estimated for 2-year cohort bins, for female cohorts born \textit{before and after} the exposure cutoff to the 1980 education reform.\footnote{Estimation by 2-year cohort bins to ensure sufficient number of observations.} Examining the magnitude and sign of the matrix’s entries across cohorts allows us to assess whether the reform affected the strength of complementarities and trade-offs between attributes. In particular, we are interested in the sign and magnitude of the \textit{Education-Education} and the \textit{Wife Education-Husband Age} coefficients.

### 7 Estimation Results

The estimates of the affinity matrix across cohorts are presented graphically in Figure \ref{fig:affinity_matrix}.\footnote{Estimates in table format are provided in Table A.14 in the appendix.} Since we characterize each spouse by two attributes - age and education, the affinity matrix will contain four cross-partial derivatives for a given cohort bin. Each dot corresponds to such a cross-partial derivative to which 95\% confidence intervals have been added. As referenced above, attributes have been standardized (mean 0, sd 1), so that the estimates’ magnitudes are directly comparable across attributes. Larger estimates indicate a greater importance of these attributes for the couple’s joint utility. Moreover, a positive coefficient indicates that two attributes are complements in generating joint utility, whereas a negative estimate indicates that the two attributes are substitutes. The affinity matrix has been separately estimated for 2-year bins based on the year of birth of the \textit{wife}.

Figure \ref{fig:affinity_matrix} reveals that education is an important attribute in the marriage market in Zimbabwe.

\footnote{Female cohorts too old to have benefited from the reform are taken to be: 1955-56; 1957-58; 1959-60; 1961-62; 1963-64. Female cohorts still young enough for the reform are: 1966-67; 1970-71; 1972-73; 1974-75; 1976-77; 1978-79. Estimation by 2-year cohort bins to ensure sufficient number of observations.}
Across all cohorts, the cross-partial Education-Education is estimated to lie between 0.4-0.5. This magnitude is roughly in line with that of the coefficient estimated by Dupuy and Galichon (2014) for their Dutch data, who find that increasing the education of spouses by 1 standard deviation increases the couple’s joint utility by 0.56 units. This complementarity gives rise to positive assortative mating on education, that is, men and women match according to their schooling levels. A comparison across birth cohorts indicates that the strength of complementarity in education appears to have been rising prior to the education reform but if anything, appears to have been slightly decreasing after the reform. A similar picture is obtained when we look at the simple correlation between years of schooling of husband and of wife (see Figure A.20). Thus, the 1980

\footnote{For instance, a larger return on investment in children’s human capital could give rise to stronger complementarities between parents’ education, as suggested by Chiappori et al. (2017).}
reform does not seem to have reinforced marital sorting on education.

In contrast to education, age by itself does not appear to be an important attribute in the marriage market in Zimbabwe. Figure 9 shows no sign of marital sorting on age and the education reform does not seem to have changed that. The cross-partial derivative $Age-Age$ is statistically indistinguishable from zero and very close to zero in magnitude across all birth cohorts of wives. The simple correlation coefficient is also very close to zero and relatively flat across cohorts (see Figure A.21).

There is suggestive evidence that, while the 1980 reform did not change the strength of complementarities on education, it might have affected trade-offs between attributes. In particular, the estimated cross-partial $Wife\ Education-Husband\ Age$ is negative throughout, indicating that these attributes are substitutes in generating match utility, but only statistically different from zero for cohorts born in 1966-69. That is, the substitutability between wife’s education and husband’s age appears to have increased for the first female cohorts affected by the reform. In these cohorts, educated wives now generate higher match utility with younger husbands, or uneducated wives generate higher match utility with older husbands. However, this substitutability fades for cohorts of wives born in 1970 or later, five years after the first cohort was exposed to the education reform. A similar transition pattern emerges when we consider the spousal age difference instead of the absolute age of each spouse, which turns the affinity matrix into a $2 \times 1$ matrix. There is a stronger substitutability between the spousal age difference and education directly after the reform comes into place, i.e. educated women initially generate higher match utility with husbands that are more similar in age (see Figure A.23).

Finally, there is some evidence that the education created transitory asymmetries in the affinity matrix, indicating that males and females temporarily do not share the same preferences for attributes. While the cross-partial $Wife\ Education-Husband\ Age$ turns negative for the first cohorts of females affected by the reform, $Wife\ Age-Husband\ Education$ remains statistically indistinguishable from zero throughout. This suggests that, while directly after the reform, educated women generated higher match utility with younger husbands, the impact of an increase in male education on joint utility was independent of their wives’ age.

---

60 Increasing wife’s education increases the couple’s joint utility when husband’s age moves in the opposite direction.
61 This dynamic is also confirmed by the simple correlation between the spousal schooling difference and spousal age difference, which turns more negative at the cutoff and then reverts to its old levels over time (see Figure A.22).
8 Discussion

Matching theories suggest that marriage outcomes depend both on preferences over a partner’s attributes and on the availability of men and women with these characteristics. In Zimbabwe, the 1980 education reform raised levels of female education and thereby altered the relative supply of educated women in the marriage market. Thus, this reform could have impacted matching patterns both by affecting own educational attainment and by shifting the education distribution of others. While the regression discontinuity design assesses the 1980 education reform’s effect on marriage outcomes in terms of spousal characteristics, the estimation of the affinity matrix allows us to examine the reform’s implications for marital sorting more carefully.

Both preference and supply forces could drive the decline in the schooling difference between spouses presented in Section 4.4. However, examining the cross-partial Education-Education in the estimated affinity matrix across cohorts indicates that the 1980 reform did not reinforce marital sorting on education. The estimates show that, while education is an important attribute for sorting, the reform does not appear to have reinforced positive assortative mating on education in Zimbabwe. Thus, the observed decline in the spousal schooling difference is attributed to supply forces, i.e. to a lack of educated men available for marriage. This is consistent with some suggestive empirical evidence indicating that a (very short-lived) increase in couples in which the wife is the more educated spouse has been driving the decline in schooling difference. Hence, it seems that the 1980 education reform initially caused educated men to be in short supply, forcing some females to marry males with fewer years of schooling, which in turn, as observed in the data, led to a lower spousal schooling difference.

In addition, both preference and supply forces could explain why educated husbands tend to come from younger cohorts after the reform. Examining the cross-partial Wife Education-Husband Age inside the estimated affinity matrix reveals that, for the first cohorts of females exposed to it, the 1980 education reform increased the substitutability between wife’s education and husband’s age. That is, educated women now generate higher utility with younger husbands. Equivalently, this could also be a sign that the reform caused older, educated men to be in short supply and thereby raised their potential for utility extraction and lowered incentives for females to marry such men. Hence, taken together, the results imply that educated women are induced to marry educated,
younger men by match utility considerations as opposed to being forced into such marriages by supply constraints. In sum, it appears that supply effects forced some females to be the more educated spouse in the marriage, while some educated females willingly married educated, younger husbands.

This empirical approach, however, also suffers from weaknesses and limitations. One major weakness is that the search and meeting technology is not explicitly modeled in this framework. Individuals choose their partner from a set of “acquaintances”, which is a random subset of the entire population of the opposite sex. This framework therefore fails to account for meeting probabilities that depend on types, that is, that a woman might be more likely to meet a college educated than a non-college educated partner if she goes to college herself. If the 1980 education reform caused men and women to spend more time together in school, educated women could be more likely to marry educated men simply because this type of man is over-represented in their consideration set. This change in meeting technology could be mistaken for a change in preferences and affect our estimated complementarities between attributes of spouses. In this context, three elements alleviate this concern. First, the affinity matrix is estimated separately by 2-year cohort bins. While we still assume a random subset of the population in a given cohort bin, we allow the population to change across cohorts. Second, there is evidence that, even after the reform, approximately 3 years elapse between the completion of schooling and start of marriage for females with secondary education. This time delay suggests that school is not the only place to meet a spouse. Third, our estimates of the affinity matrix show no spike in the Education-Education cross-partial for cohorts exposed to the 1980 education reform. If we mistake an increased meeting probability for an increase in complementarities, our Education-Education estimate should be biased upwards.

A second limitation consists in limited data availability. Ideally, we would like to characterize individuals by more attributes than just age and education. Candidate attributes that come to mind are measures of financial, physical, and personal attractiveness. However, unfortunately data on indicators of material well-being, such as income or wealth, is only available at the time of the survey, but not for the time of marriage. DHS surveys in Zimbabwe do not collect data on personality traits and only started to measure height and weight of household members from 2010 onward. This leaves us with too few observations to the left of the 1965 exposure cutoff to assess how
the education reform affected sorting on these dimensions. However, while Dupuy and Galichon (2014) find that personality traits matter, mainly through their interaction with other observable attributes, they also conclude that education is the single most important attribute in the marriage market.

A third limitation is that we do not explicitly model individuals’ schooling decision. By using a regression discontinuity design, we adopt a policy evaluation perspective and effectively assume that changes in the educational attainment are due to the reduction in the cost of schooling caused by the reform. The framework based on Dupuy and Galichon (2014) takes the distribution of education and other observed attributes of spouses as exogenously given. An equilibrium model of schooling investment and marriage decision a la Chiappori et al. (2017), however, might be an extension worth considering.

Beyond that, the large scale of the 1980 education reform in Zimbabwe might raise concerns about external validity. Yet, while only a few countries might implement such a large, nationwide reform, school expansion programs often have large local impacts. If internal migration is limited, shocking a district or a country could yield similar effects and conclusions.

9 Conclusion

This paper shows that an education reform can impact spousal characteristics and lead to trade-offs in the marriage market, while not reinforcing sorting on education. A large, nationwide education reform led to a dramatic increase in secondary school enrollment in Zimbabwe, a country in Africa where marriages were characterized by a large age difference between spouses. Zimbabwe therefore provides us with a unique setting that allows us to isolate the impact an increase in female education had on matching patterns and to examine how the marriage market adjusted to this shock over time.

Exploiting the age-specific exposure to the education reform in a regression discontinuity design, I find that females who were still young enough for secondary school at the time of the reform discontinuously acquire more schooling. These “young enough” females then marry husbands with discontinuously more schooling and there is evidence that the age difference with most of these educated husbands decreased. Females’ own schooling, however, increases by more than their hus-
bands’, so that the schooling difference between spouses actually declines. A framework that lends itself to the estimation of a so-called affinity matrix is then introduced to examine whether the observed changes in spousal characteristics were associated with a change in complementarities and marital sorting patterns. Estimating this affinity matrix both for cohorts “too old” and “young enough” reveals that, while education is an important attribute in the Zimbabwean marriage market, sorting on education was not reinforced by the reform. In part, the decline in schooling difference can be explained by a shortage of educated men forcing some females to be the more educated spouse in the marriage, implying that some educated females willingly married educated, younger males.

This paper helps to deepen our understanding of how school expansion programs can impact marriage market outcomes. While such programs have taken place in numerous developing countries, our understanding of their impact on matching patterns and marital sorting is still limited. Yet, better understanding marriage patterns is particularly important in developing countries, where - due to more limited labor market opportunities - females’ welfare largely depends on their marriage outcomes. Policy makers, before implementing such education programs, should be informed both about their direct impact on schooling and labor market outcomes and about their possible indirect impact on marriage market outcomes.
10 References


Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque (2012): “Fat-


44


Appendix

Figure A.1: Labor Market Motive for Sorting on Education

Source: Zimbabwe DHS 1988, 1994, 1999. Sample of ever married women 20-49 born 1950-1964 who were therefore too old to have benefited from the 1980 education reform. Left: probability of both spouses doing skilled work conditional on education of spouses. First index denotes husband, second index denotes wife. L: spouse has completed up to primary education, H: spouse has started secondary education. Right: estimated coefficients together with 95% confidence intervals from linear regression with both skilled (0/1) as the outcome variable and wife educated (0/1), husband educated (0/1) and both educated (0/1) as the explanatory variables. Both uneducated (0/1) is the omitted category. Region fixed effects, survey wave fixed effects, urban (0/1) and children below the age of 5 (0/1) have been added as controls. Standard errors have been clustered by age at time of the reform. Educated equals 1 if spouse has started secondary education. The probability of skilled work is increasing with the education of both spouses. Moreover, education of wives and husbands not only matter individually, but also interact with each other. This suggests complementarity in spouses’ education in labor market production as a potential motive for marital sorting on education.
Table A.1: Complementarities in Education and Female Labor Market Participation

<table>
<thead>
<tr>
<th></th>
<th>Currently working (0/1)</th>
<th>Non-agricultural work (0/1)</th>
<th>Works all year (0/1)</th>
<th>Self-employed (0/1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wife educated (0/1)</td>
<td>0.077*</td>
<td>0.164***</td>
<td>0.028</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.018)</td>
<td>(0.047)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>Husband educated (0/1)</td>
<td>0.059***</td>
<td>0.067**</td>
<td>0.064</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.028)</td>
<td>(0.040)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Both educated (0/1)</td>
<td>0.070</td>
<td>-0.036</td>
<td>0.118*</td>
<td>-0.258***</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.033)</td>
<td>(0.056)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Mean both uneducated</td>
<td>0.493</td>
<td>0.431</td>
<td>0.360</td>
<td>0.667</td>
</tr>
</tbody>
</table>

Source: Zimbabwe DHS 1988, 1994, 1999. Sample of ever married women aged 20-49 born 1950-1964 who were therefore too old to have benefited from the 1980 education reform. Educated equals one if spouse started secondary school. Estimation via linear regression controlling for region fixed effects, survey wave fixed effects, urban (0/1), and children below the age of 5 (0/1). Both uneducated (0/1) is the omitted category. Standard errors are clustered by age at time of the reform. A significant coefficient on Both educated (0/1) suggests that the interaction between spouses’ education matters beyond the contribution of their individual educational attainments. This suggests that labor market production might represent a motive for marital sorting on education in this setting.
Figure A.2: Marriages Prior to the 1980 Education Reform

Source: Zimbabwe DHS. Sample of ever married women 20-49. Averages by year of birth of wife. Ex. average years of schooling of husbands whose wives were born in 1950. Females born before 1965 were too old to have benefited from the 1980 education reform. Spousal age difference is defined as husband’s age - wife’s age. Husband’s years of schooling - wife’s years of schooling corresponds to the spousal schooling difference.
<table>
<thead>
<tr>
<th>Cohorts</th>
<th>1950-64</th>
<th>1966-80</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>Age at survey</td>
<td>7,060</td>
<td>37.8</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>7,060</td>
<td>5.2</td>
</tr>
<tr>
<td>Started secondary (0/1)</td>
<td>7,060</td>
<td>0.19</td>
</tr>
<tr>
<td>Never married (0/1)</td>
<td>6,194</td>
<td>0.02</td>
</tr>
<tr>
<td>Currently married (0/1)</td>
<td>6,194</td>
<td>0.75</td>
</tr>
<tr>
<td>Age first marriage</td>
<td>6,067</td>
<td>18.9</td>
</tr>
<tr>
<td>Spousal age difference</td>
<td>2,280</td>
<td>8.0</td>
</tr>
<tr>
<td>Husband’s years of schooling</td>
<td>5,822</td>
<td>6.5</td>
</tr>
<tr>
<td>Spousal schooling difference</td>
<td>5,822</td>
<td>1.4</td>
</tr>
<tr>
<td>Worked last week (0/1)</td>
<td>7,056</td>
<td>0.51</td>
</tr>
<tr>
<td>Skilled work (0/1)</td>
<td>3,098</td>
<td>0.34</td>
</tr>
<tr>
<td>Age first birth</td>
<td>6,031</td>
<td>19.4</td>
</tr>
<tr>
<td>Total number of kids</td>
<td>6,194</td>
<td>4.6</td>
</tr>
<tr>
<td>Ideal number of kids</td>
<td>5,959</td>
<td>5.3</td>
</tr>
</tbody>
</table>

Source: Zimbabwe DHS. Sample of women aged 30-49 for marriage and fertility outcomes. Females born before 1965 were too old to have benefited from the 1980 education reform, while females born after 1965 were still young enough to benefit. Spousal age difference = husband’s age - wife’s age. Spousal schooling difference = husband’s years of schooling - wife’s years of schooling.
Figure A.3: Spousal Differences Prior to the 1980 Education Reform

Source: Zimbabwe DHS. Sample of ever married women 20-49 born 1950-1964. Histograms. Females born before 1965 were too old to have benefited from the 1980 education reform. Spousal age difference = husband’s age - wife’s age. Spousal schooling difference = husband’s years of schooling - wife’s years of schooling.
Figure A.4: **Spousal Age Differences in Southern and Eastern Africa**

![Kernel Density](image)

Source: DHS. Malawi 2000, Mozambique 1997, Tanzania 2004, Zambia 1996, Zimbabwe 1999. Sample of women in union, aged 15-49, whose husband’s age is known. Kernel density. Spousal age difference = husband’s age - wife’s age. The age difference distributions are skewed to the right, have a concentration around the mode, and a low representation of negative values.

Figure A.5: **Gross Secondary School Enrollment in Zimbabwe**

![Secondary Enrollments](image)

Source: UNESCO Institute for Statistics. Total enrollment (number of students) in all secondary schools. Between 1979 and 1985, secondary school enrollment increased from 66,215 to 482,000 (+628%). As a comparison, total enrollment increased from 885,801 to 2,698,878 (+628%).
Figure A.6: Transition Rates from Primary to Secondary Schools in Zimbabwe

![Transition Rates from Primary to Secondary Schools in Zimbabwe](image)

Source: Agüero and Bharadwaj (2014). Transition rate is defined as the number of students in the first grade of secondary school divided by the number of students in the last grade of primary school. The transition rate from primary to secondary school increased from 27% in 1979 to over 80% in 1980.

Figure A.7: Number of Primary and Secondary Schools in Zimbabwe

![Number of Primary and Secondary Schools in Zimbabwe](image)

Source: Nhundu (1992). Trends in school construction by education level, 1979–88. Elaborated based on data from Zimbabwe Ministry of Education, Annual reports of the Secretary for Education, various years. Between 1979 and 1985, the number of primary schools increased by 76% while the number of secondary schools increased by 586%.
Figure A.8: Teachers in Primary and Secondary Schools in Zimbabwe

Source: World Bank Development Indicators. Number of teachers and student-teacher ratio in primary and secondary education. The number of teachers rose to accommodate the increase in schooling demand while the student-teacher ratio remained at reasonable levels.
Figure A.9: Proportions of Females with Any Schooling

Source: Zimbabwe DHS 1988, 1994, 1999, 2005, 2010, 2015. Sample of women 20-49. Proportions of females who have at least one year of schooling by year of birth. Vertical line drawn at cutoff birth cohort of 1965. Females born before 1965 were too old to have benefited from the 1980 education reform, while females born after 1965 were still young enough. Females born after 1965 are more likely to have had at least one year of schooling.

Figure A.10: Proportions of Females by Education Level

Source: Zimbabwe DHS 1988, 1994, 1999, 2005, 2010, 2015. Sample of women 20-49. Proportions of females who started a level of schooling by year of birth. Primary corresponds to up to 7 years of schooling, junior secondary to 8-9 years of schooling, middle secondary to 10-11 years of schooling, and advanced secondary to 12-13 years of schooling. Vertical line drawn at cutoff birth cohort of 1965. Females born before 1965 were too old to have benefited from the 1980 education reform, while females born after 1965 were still young enough. The relevant margin of this reform is going and completing up to 4 years of secondary school.
Figure A.11: Education Distribution of Females in Zimbabwe

Figure A.12: Robustness: First Stage of Males

Source: Zimbabwe Census 2012. Sample of men 32-60. Averages by year of birth. Vertical line drawn at cutoff birth cohort of 1965. Males born before 1965 were too old to benefit from the 1980 education reform, while males born after 1965 were still young enough. No clear point of discontinuity for males.
Figure A.13: Reduced Form Females: Age at First Marriage

Figure A.14: Comparing Education of Spouses

Source: Zimbabwe DHS 1988, 1994, 1999, 2005, 2010, 2015. Sample of ever married women 20-49. Comparison based on years of schooling of husband and wife. A couple counts as *Wife More Educated* when the wife has more years of schooling than her husband. Proportions by year of birth of wife. Vertical line drawn at cutoff birth cohort of 1965. Females born before 1965 were too old to benefit from the 1980 education reform, while females born after 1965 were still young enough. Females born after 1965 increasingly more likely to be the more educated spouse in the marriage for the first few cohorts.
Figure A.15: Proportion of Couples by Education of Spouses

Source: Zimbabwe DHS 1988, 1994, 1999, 2005, 2010, 2015. Sample of ever married women 20-49. Left: Proportion of couples in which only spouse had started secondary school. Right: Proportion of couples in which either none of the spouses or both spouses had started secondary school. Proportions by year of birth of wife. Vertical line drawn at cutoff birth cohort of 1965. Females born before 1965 were too old to benefit from the 1980 education reform, while females born after 1965 were still young enough. The first few cohorts of females born after 1965 who have some secondary education are more likely to be married to a husband who did not go to secondary school.
Figure A.16: Probability Never Marry for Males

By Education Level

Figure A.17: Reduced Form: Spousal Age Difference

Source: Zimbabwe DHS 1999, 2005, 2010, 2015. Sample of ever married women 20-49. Left: Spousal age difference = husband’s age - wife’s age. Averages by year of birth of wife. Ex. average spousal age difference for couples where the wife is born in 1950. Right: Proportion of couples where the husband is at least six years older. Vertical line drawn at cutoff birth cohort of 1965. Point estimates and standard errors of nonparametric 2SLS regressions have been added. Females born before 1965 were too old to benefit from the 1980 education reform, while females born after 1965 were still young enough. No clear evidence of a discontinuity in average age difference, but decline in couples with large age difference.
Figure A.18: Distribution of Spousal Age Difference Before and After Reform

Table A.3: Spousal Age Difference

<table>
<thead>
<tr>
<th>Spousal Age Difference up to</th>
<th>2SLS</th>
<th>First Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Years</td>
<td>0.050**</td>
<td>1.598***</td>
</tr>
<tr>
<td>3 Years</td>
<td>0.052**</td>
<td></td>
</tr>
<tr>
<td>4 Years</td>
<td>0.056**</td>
<td></td>
</tr>
<tr>
<td>5 Years</td>
<td>0.046*</td>
<td></td>
</tr>
<tr>
<td>6 Years</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>7 Years</td>
<td>-0.005</td>
<td></td>
</tr>
</tbody>
</table>

Note: Females born in 1965 were 15 years old at the time of the education reform and 1980 and therefore still young enough for secondary school to benefit. Nonparametric 2SLS estimation. Each column corresponds to a separate regression with an indicator for a couple belonging to a given age difference (0/1) as the outcome variable. Females young enough to benefit from the reform are more likely to marry husbands more similar in age.
Figure A.19: Proportion of Husbands with Secondary Schooling by Spousal Age Difference

Table A.4: Nonparametric Estimation Results

<table>
<thead>
<tr>
<th></th>
<th>Never Married</th>
<th>Husband’s yrs of schooling</th>
<th>Spousal school difference</th>
<th>Spousal age difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2SLS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>0.002</td>
<td>0.882***</td>
<td>-0.378***</td>
<td>-0.418</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.232)</td>
<td>(0.127)</td>
<td>(0.423)</td>
</tr>
<tr>
<td><strong>First Stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young enough in 1980</td>
<td>1.313***</td>
<td>1.151***</td>
<td>1.380***</td>
<td>1.323***</td>
</tr>
<tr>
<td></td>
<td>(0.346)</td>
<td>(0.324)</td>
<td>(0.230)</td>
<td>(0.382)</td>
</tr>
</tbody>
</table>

Source: Zimbabwe DHS. Sample of women 25-49 for Never Married and sample of ever married women 20-49 otherwise. Spousal schooling difference = husband’s years of schooling - wife’s years of schooling. Spousal age difference = husband’s age - wife’s age. Females born in 1965 were 15 years old at the time of the education reform and 1980 and therefore still young enough for secondary school to benefit. Baseline estimation: MSE-optimal bandwidth selector with distinct bandwidths on each side of cutoff, survey weights and heteroskedasticity-robust standard errors computed using nearest neighbor methods.
Table A.5: Nonparametric Robustness: Husband’s Years of Schooling

<table>
<thead>
<tr>
<th></th>
<th>Outcome: Husband’s Years of Schooling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2SLS</td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>0.882*** 0.774** 1.036*** 0.684 0.677*** 0.921*** 0.882*** 0.803*** 0.850***</td>
</tr>
<tr>
<td></td>
<td>(0.232) (0.331) (0.330) (0.684) (0.215) (0.287) (0.231) (0.175) (0.196)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>First Stage: Years of Schooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Enough in 1980</td>
<td>1.151*** 1.130** 1.017*** 0.749 1.127*** 0.969*** 1.151*** 1.251** 1.155</td>
</tr>
<tr>
<td></td>
<td>(0.324) (0.469) (0.390) (0.539) (0.292) (0.302) (0.324) (0.613) (0.772)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bandwidth Selection</th>
<th>MSE</th>
<th>CER</th>
<th>MSE</th>
<th>CER</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Common</td>
<td>Common</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
</tr>
<tr>
<td>Weights</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Urban</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Variance</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>HC0</td>
<td>NN</td>
<td>NN</td>
</tr>
<tr>
<td>SE Robust</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>SE Clustered</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Region</td>
<td>Regions</td>
<td>Age80</td>
</tr>
</tbody>
</table>

Table A.6: **Nonparametric Robustness: Spousal Schooling Difference**

**Outcome: Spousal Schooling Difference**

### 2SLS

<table>
<thead>
<tr>
<th>Years of Schooling</th>
<th>-0.378***</th>
<th>-0.325*</th>
<th>-0.270</th>
<th>-0.206</th>
<th>-0.360***</th>
<th>-0.325*</th>
<th>-0.315**</th>
<th>-0.313**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.127)</td>
<td>(0.188)</td>
<td>(0.167)</td>
<td>(0.217)</td>
<td>(0.117)</td>
<td>(0.182)</td>
<td>(0.141)</td>
<td>(0.138)</td>
</tr>
</tbody>
</table>

### First Stage: Years of Schooling

<table>
<thead>
<tr>
<th>Young Enough in 1980</th>
<th>1.380***</th>
<th>1.767***</th>
<th>1.352***</th>
<th>1.514***</th>
<th>1.679***</th>
<th>1.229***</th>
<th>1.531***</th>
<th>1.617**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.230)</td>
<td>(0.449)</td>
<td>(0.303)</td>
<td>(0.436)</td>
<td>(0.262)</td>
<td>(0.270)</td>
<td>(0.291)</td>
<td>(0.660)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bandwidth Selection</th>
<th>MSE</th>
<th>CER</th>
<th>MSE</th>
<th>CER</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weights</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Urban</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Variance</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>HC0</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
</tr>
<tr>
<td>SE Robust</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>SE Clustered</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Region</td>
<td>No</td>
<td>Region/Age80</td>
</tr>
</tbody>
</table>

Table A.7: **Nonparametric Robustness: Spousal Age Difference**

**Outcome: Spousal Age Difference**

| 2SLS |  
|---|---|---|---|---|---|---|---|---|---|
| Years of Schooling | -0.418 | -0.586 | -0.316 | -0.586 | -0.346 | -0.411 | -0.418 | -0.418 | -0.418 |
| (0.423) | (0.539) | (0.601) | (0.475) | (0.386) | (0.475) | (0.421) | (0.332) | (0.338) |

**First Stage: Years of Schooling**

| Young Enough in 1980 | 1.323*** | 2.129*** | 1.569** | 1.688*** | 1.355*** | 1.177*** | 1.323*** | 1.323** | 1.323** |
| (0.382) | (0.768) | (0.644) | (0.516) | (0.320) | (0.341) | (0.385) | (0.656) | (0.621) |

<table>
<thead>
<tr>
<th>Bandwidth Selection</th>
<th>MSE</th>
<th>CER</th>
<th>MSE</th>
<th>CER</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
<td>Distinct</td>
</tr>
<tr>
<td>Weights</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Urban</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Variance</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>NN</td>
<td>HC0</td>
<td>NN</td>
</tr>
<tr>
<td>Robust</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>SE</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table A.8: **Parametric Robustness: Husband’s Years of Schooling**

<table>
<thead>
<tr>
<th></th>
<th>Age 0-30 in 1980</th>
<th>Age 5-25 in 1980</th>
<th>Age 10-20 in 1980</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2SLS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>0.573***</td>
<td>0.639***</td>
<td>0.427</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.097)</td>
<td>(0.342)</td>
</tr>
<tr>
<td></td>
<td>0.576***</td>
<td>0.572***</td>
<td>0.208</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.208)</td>
<td>(0.796)</td>
</tr>
<tr>
<td></td>
<td>0.558</td>
<td>0.588***</td>
<td>0.809**</td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td>(0.376)</td>
<td>(0.376)</td>
</tr>
<tr>
<td></td>
<td>0.588***</td>
<td>0.558</td>
<td>0.280</td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td>(0.376)</td>
<td>(0.376)</td>
</tr>
<tr>
<td></td>
<td>0.797***</td>
<td>0.797***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.280)</td>
<td>(0.280)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Stage: Years of Schooling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young Enough in 1980</td>
<td>1.896***</td>
<td>1.297***</td>
<td>0.578*</td>
</tr>
<tr>
<td></td>
<td>(0.144)</td>
<td>(0.226)</td>
<td>(0.300)</td>
</tr>
<tr>
<td></td>
<td>1.624***</td>
<td>0.846***</td>
<td>0.296</td>
</tr>
<tr>
<td></td>
<td>(0.169)</td>
<td>(0.251)</td>
<td>(0.367)</td>
</tr>
<tr>
<td></td>
<td>0.937***</td>
<td>0.718**</td>
<td>0.718**</td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td>(0.323)</td>
<td>(1.032)</td>
</tr>
<tr>
<td></td>
<td>1.985*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.032)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>17,975</td>
<td>17,975</td>
<td>17,975</td>
</tr>
<tr>
<td>Polynomial</td>
<td>Linear</td>
<td>Quadratic</td>
<td>Cubic</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Window width (yrs)</td>
<td>15</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Window width (yrs)</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Window width (yrs)</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Interaction</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Region FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Survey FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Table A.9: **Parametric Robustness: Spousal Schooling Difference**

Outcome: Spousal Schooling Difference

<table>
<thead>
<tr>
<th></th>
<th>Age 0-30 in 1980</th>
<th>Age 5-25 in 1980</th>
<th>Age 10-20 in 1980</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2SLS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>-0.425***</td>
<td>-0.372***</td>
<td>-0.572*</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.094)</td>
<td>(0.335)</td>
</tr>
<tr>
<td>First Stage: Years of Schooling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young Enough in 1980</td>
<td>1.896***</td>
<td>1.297***</td>
<td>0.578*</td>
</tr>
<tr>
<td></td>
<td>(0.144)</td>
<td>(0.226)</td>
<td>(0.300)</td>
</tr>
<tr>
<td>Observations</td>
<td>17,975</td>
<td>17,975</td>
<td>17,975</td>
</tr>
<tr>
<td>Polynomial</td>
<td>Linear</td>
<td>Quadratic</td>
<td>Cubic</td>
</tr>
<tr>
<td>Window width (yrs)</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Interaction</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Region FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Survey FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table A.10: **Parametric Robustness: Spousal Age Difference**

**Outcome: Spousal Age Difference**

<table>
<thead>
<tr>
<th></th>
<th>Age 0-30 in 1980</th>
<th>Age 5-25 in 1980</th>
<th>Age 10-20 in 1980</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2SLS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>-0.121 (0.153)</td>
<td>-0.166 (0.347)</td>
<td>-0.537 (0.721)</td>
</tr>
<tr>
<td></td>
<td>-0.143 (0.216)</td>
<td>-0.405 (0.565)</td>
<td>-0.647 (0.715)</td>
</tr>
<tr>
<td></td>
<td>-0.249 (0.387)</td>
<td>-1.027 (0.723)</td>
<td>-1.638** (0.807)</td>
</tr>
<tr>
<td><strong>First Stage: Years of Schooling</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young Enough in 1980</td>
<td>1.896*** (0.144)</td>
<td>1.297*** (0.226)</td>
<td>0.578* (0.300)</td>
</tr>
<tr>
<td></td>
<td>1.624*** (0.169)</td>
<td>0.846*** (0.251)</td>
<td>0.296 (0.367)</td>
</tr>
<tr>
<td></td>
<td>0.937*** (0.200)</td>
<td>0.718** (0.323)</td>
<td>1.985* (1.032)</td>
</tr>
<tr>
<td>Observations</td>
<td>10,497</td>
<td>10,497</td>
<td>6,829</td>
</tr>
<tr>
<td>Polynomial</td>
<td>Linear</td>
<td>Quadratic</td>
<td>Cubic</td>
</tr>
<tr>
<td>Window width (yrs)</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Interaction</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Region FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Survey FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table A.11: **Robustness: Sample Ever Married Black Women** (Nonparametric Estimation)

<table>
<thead>
<tr>
<th></th>
<th>Never married</th>
<th>Husband’s yrs of schooling</th>
<th>Husband started secondary school</th>
<th>Spousal school difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2SLS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>-0.013</td>
<td>0.888***</td>
<td>0.146***</td>
<td>-0.253</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.316)</td>
<td>(0.047)</td>
<td>(0.322)</td>
</tr>
<tr>
<td><strong>First Stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young enough in 1980</td>
<td>1.227**</td>
<td>1.034***</td>
<td>1.088***</td>
<td>1.201***</td>
</tr>
<tr>
<td></td>
<td>(0.505)</td>
<td>(0.373)</td>
<td>(0.343)</td>
<td>(0.388)</td>
</tr>
</tbody>
</table>

Source: Zimbabwe DHS 1988, 1994 (ethnicity is only available in these two survey waves, where 98.2% report being Black). Sample of ever married Black women. Spousal schooling difference = husband’s years of schooling - wife’s years of schooling. Husband’s age only becomes available from DHS 1999 onward. Females born in 1965 were 15 years old at the time of the education reform and 1980 and therefore still young enough for secondary school to benefit. Baseline estimation: MSE-optimal bandwidth selector with distinct bandwidths on each side of cutoff, survey weights and heteroskedasticity-robust standard errors computed using nearest neighbor methods.
Table A.12: **Robustness: Sample Ever Married Women 25-49** (Nonparametric Estimation)

<table>
<thead>
<tr>
<th></th>
<th>Never married</th>
<th>Husband’s yrs of schooling</th>
<th>Spousal school difference</th>
<th>Spousal age difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2SLS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>0.002</td>
<td>1.109***</td>
<td>-0.417***</td>
<td>-0.418</td>
</tr>
<tr>
<td>(0.013)</td>
<td>(0.324)</td>
<td>(0.126)</td>
<td>(0.423)</td>
<td></td>
</tr>
<tr>
<td><strong>First Stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young enough in 1980</td>
<td>1.313***</td>
<td>1.138***</td>
<td>1.504***</td>
<td>1.323***</td>
</tr>
<tr>
<td>(0.346)</td>
<td>(0.428)</td>
<td>(0.254)</td>
<td>(0.382)</td>
<td></td>
</tr>
</tbody>
</table>

Source: Zimbabwe DHS. Sample of ever married women 25-49. Spousal schooling difference = husband’s years of schooling - wife’s years of schooling. Spousal age difference = husband’s age - wife’s age. Females born in 1965 were 15 years old at the time of the education reform and 1980 and therefore still young enough for secondary school to benefit. Baseline estimation: MSE-optimal bandwidth selector with distinct bandwidths on each side of cutoff, survey weights and heteroskedasticity-robust standard errors computed using nearest neighbor methods.
Table A.13: **Robustness: Sample of Monogamous Wives** (Nonparametric Estimation)

<table>
<thead>
<tr>
<th></th>
<th>Husband’s yrs of schooling</th>
<th>Spousal school difference</th>
<th>Spousal age difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SLS</td>
<td>0.675***</td>
<td>-0.363***</td>
<td>-0.260</td>
</tr>
<tr>
<td></td>
<td>(0.207)</td>
<td>(0.169)</td>
<td>(0.371)</td>
</tr>
<tr>
<td>First Stage</td>
<td>1.053***</td>
<td>1.183***</td>
<td>1.261***</td>
</tr>
<tr>
<td></td>
<td>(0.278)</td>
<td>(0.269)</td>
<td>(0.338)</td>
</tr>
</tbody>
</table>

### Table A.14: Estimation Results Affinity Matrix

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diagonal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-Age</td>
<td>0.077</td>
<td>-0.006</td>
<td>0.007</td>
<td>0.005</td>
<td>0.003</td>
<td>0.007</td>
<td>0.007</td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Education-Education</td>
<td>0.339***</td>
<td>0.380***</td>
<td>0.413***</td>
<td>0.500***</td>
<td>0.520***</td>
<td>0.547***</td>
<td>0.524***</td>
<td>0.487***</td>
<td>0.491***</td>
<td>0.437***</td>
<td>0.466***</td>
<td>0.449***</td>
</tr>
<tr>
<td><strong>Off-Diagonal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wife Age, Husband Education</td>
<td>-0.070</td>
<td>0.034</td>
<td>-0.041</td>
<td>-0.034</td>
<td>-0.056</td>
<td>-0.048</td>
<td>-0.041</td>
<td>-0.010</td>
<td>-0.014</td>
<td>-0.018</td>
<td>-0.009</td>
<td>-0.010</td>
</tr>
<tr>
<td>Wife Education, Husband Age</td>
<td>-0.085</td>
<td>-0.072</td>
<td>-0.066</td>
<td>-0.073</td>
<td>-0.028</td>
<td>-0.083*</td>
<td>-0.089**</td>
<td>-0.018</td>
<td>-0.038</td>
<td>-0.030</td>
<td>-0.015</td>
<td>-0.034</td>
</tr>
<tr>
<td><strong>Couples</strong></td>
<td>150</td>
<td>263</td>
<td>251</td>
<td>393</td>
<td>454</td>
<td>501</td>
<td>659</td>
<td>761</td>
<td>1,010</td>
<td>1,086</td>
<td>1,205</td>
<td>1,121</td>
</tr>
</tbody>
</table>

*** p<0.01; ** p<0.05; * p<0.1. Source: Zimbabwe DHS 1999, 2005, 2010, 2015. Sample of ever married monogamous couples. Estimation by 2-year bins of wives’ birth cohorts. Each entry corresponds to an entry of the $2 \times 2$ affinity matrix based on age and education and thus, is a cross-partial derivative. Attributes have been standardized, so that estimates are directly comparable. Positive entries indicate complementarity, whereas negative entries indicate substitutability between attributes. Females born before 1965 were too old to benefit from the 1980 education reform, while females born after 1965 were still young enough.

### Figure A.20: Correlation between Schooling of Husband and Schooling of Wife

Figure A.21: Correlation between Age of Husband and Age of Wife


Figure A.22: Correlation b/w Spousal Schooling Difference and Spousal Age Difference

Figure A.23: Robustness: Estimates of Entries of Affinity Matrix $A$ with Spousal Age Difference

Source: Zimbabwe DHS 1999, 2005, 2010, 2015. Sample of ever married monogamous couples. Estimation by 2-year bins of wives’ birth cohorts. Each dot corresponds to an entry of the $2 \times 1$ affinity matrix based on spousal age difference and education and thus, is a cross-partial derivative. Attributes have been standardized, so that estimates are directly comparable. Positive entries indicate complementarity, whereas negative entries indicate substitutability between attributes. 95% confidence intervals have been added and a horizontal line at 0 has been drawn. Vertical line drawn at cutoff birth cohort of 1965. Females born before 1965 were too old to benefit from the 1980 education reform, while females born after 1965 were still young enough.