Essays on Schooling and Health in Sub-Saharan Africa

ESSAYS ON SCHOOLING AND HEALTH IN SUB-SAHARAN AFRICA

JAN-WALTER DE NEVE

A Dissertation Submitted To The Faculty Of

The Harvard T.H. Chan School of Public Health

in Partial Fulfillment of the Requirements

for the Degree of *Doctor of Science*

in the Department of *Global Health and Population*

Harvard University

Boston, Massachusetts.

May 2017

Dissertation Advisor: **Till Bärnighausen** **Jan-Walter De Neve**

Essays on Schooling and Health in sub-Saharan Africa

Abstract

This dissertation includes three papers in the field of education and health.

This first paper provides estimates of the effects of secondary schooling on HIV infection and childbearing. A 1996 policy reform in Botswana changed the grade structure of secondary school and led to sharp increases in educational attainment among affected birth cohorts. Exploiting cohort-specific exposure and differential impact by birth village as 'natural experiments', I find that the reform decreased HIV infection and childbearing. Turning to mechanisms, I find that schooling had no effect on HIV knowledge; however it influenced sexual debut, norms and behaviors, labor force participation, and literacy. In HIV endemic settings, estimates of the returns to secondary schooling may be underestimated due to the exclusion of health benefits.

The second paper uses data from the Tanzania Censuses to investigate whether additional schooling in children affects their parent's old-age survival. To do so, I exploit quasi-random variation resulting from a policy reform that initiated a drive towards Universal Primary Education in the mid-1970s. I show that the reform caused a large increase in school enrollment. Additional primary schooling in children as a result of the reform induced large reductions in the probability of parental death by the time exposed child cohorts had reached age 40. These survival gains highlight the large societal benefits of human capital investment, particularly in low-resource settings.

In the third paper, I exploit the 1980 Zimbabwe education reform, which rapidly expanded access to secondary schools. Exploiting cohort-specific exposure, I assess whether additional schooling among parents affected the risk of undernutrition in their children under five. I find no evidence of a causal effect of parental schooling on undernutrition in their children at the national level. Among urban and wealthier households, additional maternal schooling had a protective effect against the risk of wasting in their children. Parental schooling may play a more muted role in child investment decisions than previously suggested. This essay emphasizes the need for directed health investments to improve child development in low-resource settings.

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I am indebted to my research committee: Till Bärnighausen, Günther Fink, and Subu Subramanian. I also thank Jacob Bor, my coauthor on the first paper, for his mentorship and friendship throughout my doctoral studies. I also thank Max Essex and Sikhulile Moyo for connecting me to the Botswana Harvard AIDS Institute.

Many others have provided valuable comments and feedback on the essays in this dissertation, including Richard Tabulawa, Happy Siphambe, Michael R. Reich, Brothers Malema, Lawrence Katz, Pilate Khulumani, as well as seminar participants at Harvard University, Botswana Harvard AIDS Institute, Harvard Population and Development Center, Boston University School of Public Health, Oxford University, PAA, PopPov, and NEUDC. I also thank Ria Madison, Bernadette Kgakge, Virginia Sebekedi, Thapelo Maruatona, Nametso Kgosiyame, Koontse Mokgwathi, and Susan Matroos for assistance with obtaining data, and the respondents and staff of the Botswana Census, AIDS Impact Surveys, Labor Force Surveys, Tanzania Population Censuses, and the Demographic and Health Surveys of Zimbabwe.

For administrative and financial support, I am grateful to Barbara Heil, Allison Gallant, Amy Levin, Efren Mencia, as well as to the Taiplin Fellowship and Takemi Program in International Health at the Harvard T.H. Chan School of Public Health, Belgian American Educational Foundation, Fernand Lazard Foundation, and the Boston University School of Public Health. Finally, I thank the Harvard-MIT Data Center for providing computing facilities and infrastructure in a collaborative work environment. The sponsors had no role in study design, data collection, analysis, and interpretation, or writing of the essays. Any omissions or errors are my own.

To Pınar and Aren

Introduction

This dissertation explores three aspects of the relationship between schooling and health.

First, education and health are strongly correlated, but there is no consensus on whether more schooling is a causal determinant of better health (Clark and Royer 2013, Cutler and Lleras-Muney 2006, 2010, Grossman 2015, Lleras-Muney 2005, Ross and Wu 1995). Education and health are both aspects of human capital that enable people to earn more (Bor et al. 2012, Mincer 1974) and engender people's capabilities (Sen 1980) to lead fulfilling lives. How these different aspects of human capital are co-produced has become a topic of recent economic research (Case, Lubotsky, and Paxson 2002, Currie 2009, Currie and Moretti 2003). Understanding the causal structure can point to opportunities for high return investments. For example, some have argued that investments in education would have a larger impact than investments in medical care (Galea et al. 2011). If the relationship is causal, then expanding schooling opportunities would help reduce the wide disparities in health across populations (Berkman and Kawachi 2000, Cutler, Deaton, and Lleras-Muney 2006, Glied and Lleras-Muney 2008, Kaufman and Cooper 1999, Krieger 2001, Meara, Richards, and Cutler 2008). This dissertation explores some of the largest schooling reforms in recent history to generate estimates of the causal effects of education on health.

Second, in addition to the question of whether education causally improves health, it is unclear which levels of schooling are most important. The Heckman-Cunha model of dynamic complementarities points to early childhood as the best time to invest, because higher human capital at younger ages increases future acquisition of human capital (Cunha and Heckman 2007). Previous investments in early childhood may function as multipliers for future investments. In the first paper of this dissertation, I consider the returns to human capital investment in another potentially "critical period": late adolescence, in particular the years of upper secondary school. Late adolescence matters for several reasons: (1) the development of new skills (e.g., complex reasoning and future-oriented thinking) with strong returns in the labor market (Crews, He, and Hodge 2007, Dahl 2004, Steinberg 2005); (2) many adolescents are making their own decisions for the first time; and (3) the decisions facing adolescents are important and have significant path dependence, e.g. pregnancy (Dinçer, Kaushal, and Grossman 2014, Grépin and Bharadwaj 2015, Osili and Long 2008), and planning around education and a career path (Wilde, Batchelder, and Ellwood 2010). Dynamic complementarities from labor market experience and (possibly) home production (Wilde, Batchelder, and Ellwood 2010) imply strong returns to specialization in late adolescence and large impacts of decisions in late adolescence. Late adolescence may function as a "railroad switch", in which individuals sort into career and non-career tracks, leading to divergence and clustering of multiple behaviors and outcomes.

Third, relatively little is known about the intergenerational transmission of human capital (De Neve and Kawachi 2017). In papers two and three of this dissertation, I assess the consequences of additional human capital in children for the health of both the younger and older generation. More highly educated children may serve as a substitute for market or government institutions for family members (Lillard and Willis 1997, Mason and Lee 2006); better communicate health knowledge and skills acquired at or after school, positively affecting the health behavior of family members (Berkman et al. 2000, Cutler and Lleras-Muney 2010, Field and de la Roca 2005, Rowa-Dewar, Amos, and Cunningham-Burley 2014); be exposed to a wide range of beliefs and norms at school, such as around norms against domestic violence (Rende 2014), or making them more familiar with "modern society" that may make them more receptive to modern medicine (Aslam and Kingdon 2012, Frost, Forste, and Haas 2005); and they may opt to pursue a career in the health sector, allowing the provision of direct care to other household members (Bauman et al. 2006, Evans and Becker 2009, McGarry 1998).

Differential investment in children's human capital could thus be a key driver of health inequalities among their core—and possibly extended—family members (Berkman and Kawachi 2000, Kaufman and Cooper 1999, Krieger 2001, Marmot 2003).

Finally, I examine the relationship between schooling and health in a particular context: that of a large, generalized HIV epidemic and limited public welfare.

Chapter 1

A Social Vaccine? HIV Infection, Childbearing and the Non-Pecuniary Returns to Secondary Schooling in Botswana[1]

1.1 Introduction

HIV remains a major global health challenge, with an estimated 2.1 Million new infections each year (Joint United Nations Programme on HIV/AIDS 2013). Formal education, particularly of young women, has been proposed as a "social vaccine" to reduce the spread of HIV (Vandemoortele and Delamonica 2000). However, there is little causal evidence to support this claim (Jukes, Simmons, and Bundy 2008). Early national surveillance surveys found higher rates of HIV among people with more education in a number of sub-Saharan Africa countries (Dallabetta et al. 1993, Fortson 2008, Kirunga and Ntozi 1997). However, more recent studies have found a protective association between higher education and HIV infection, as information on HIV prevention strategies has become more widely available and the epidemic has matured (Barnighausen et al. 2007, Fylkesnes et al. 2001, Hargreaves et al. 2008, Iorio and Santaeulalia-Llopis 2011).

The effect of education on HIV risk is theoretically ambiguous. Education may reduce HIV risk through: increased exposure to information about HIV and prevention methods (Agüero and Bharadwaj 2014, de Walque 2009); improved cognitive skills and ability to navigate a complex risk environment (Cutler and Lleras-Muney 2010); higher returns to market labor (Psacharopoulos 1994, Schultz 1960, Siphambe 2000, Willis

[1]Co-authored with Jacob Bor. This article expands on (De Neve et al. 2015) —a research paper that reports on the causal effect of secondary school on HIV status in the Botswana AIDS Impact Surveys II (2004) and III (2008) samples.

1973), increasing financial independence of women, reducing participation in transactional sex, and increasing bargaining power within relationships (Baird et al. 2012, Dunkle et al. 2004, Mabsout and van Staveren 2010); assortative mating with lower-risk partners (Baird et al. 2012, Epstein and Guttman 1984, Pettifor et al. 2005); less time for sexual relationships leading to increased abstinence during years in school (Alsan and Cutler 2013, Black, Devereux, and Salvanes 2008); changing childbearing preferences, leading to increased contraceptive use and/or less frequent sex (Kravdal 2002, Martin 1995, McCrary and Royer 2011, Monstad, Propper, and Salvanes 2008, Osili and Long 2008, Rosenberg et al. 2015) and increased future orientation. On the other hand, education may increase the size of one's sexual network; prolong the period of pre-marital sex (Case and Paxson 2013); increase earnings, enabling men to have more partners and/or engage in riskier sex (Kohler and Thornton 2011); and increase one's attractiveness in the "market" for sexual partners, leading (possibly) to more opportunities for unprotected sex.

This paper contributes to a larger literature on the health impacts of education in general. Due to the paucity of randomized trials and natural experiments, there is little consensus in the economics and public health literature on whether a causal relationship exists between education and health, and if so what the mechanisms are health (Clark and Royer 2013, Cutler and Lleras-Muney 2006, 2010, Grossman 2015, Lleras-Muney 2005, Ross and Wu 1995). In addition to its significance for HIV prevention, this study thus has implications for understanding the role of education in the health production function, and its effect on disparities in health outcomes across populations (Berkman and Kawachi 2000, Cutler, Deaton, and Lleras-Muney 2006, Glied and Lleras-Muney 2008, Kaufman and Cooper 1999, Krieger 2001, Meara, Richards, and Cutler 2008).

The challenge in determining the causal effect of schooling on HIV infection risk is that school attainment is closely related to factors such as socioeconomic status, psychological traits, and preferences, which are difficult to control for fully in observational studies and which may also affect HIV risk. Several randomized trials have identified protective effects of school supports on HIV risk factors (Cho et al. 2011) and correlated outcomes (teen pregnancy) (Duflo, Dupas, and Kremer 2015), but none has shown an effect on HIV infection itself. A conditional cash transfer to keep girls in school reduced HIV risk in a Malawi study, but the effect appeared to be due to the cash not the condition (Baird et al. 2012). A challenge for schooling RCTs is that the demand-side subsidies offered have generally led to small changes in educational attainment, with low power to identify an effect on HIV infection. In contrast, large supply-side policy reforms have, in some cases, led to quite large changes in educational attainment (Borkum 2009, Chicoine 2012, Duflo 2001), and thus suggest a promising approach to assess the causal effect of schooling on HIV infection risk.

In 1996, the Government of Botswana changed the grade structure of secondary school nationwide, moving grade ten from senior secondary school to junior secondary school, dramatically increasing the number of seats available. The result was a sharp increase in average years of schooling by 0.8 years at the population level. The policy change affected specific birth cohorts – e.g., those who would have entered junior secondary school in 1996 or later – and as a sector-specific supply-side reform was unlikely to have affected outcomes through mechanisms other than schooling itself. I use cohort variation in exposure to the policy reform to identify the effect of secondary education on HIV status, childbearing, and labor market outcomes. This paper builds on recently-published work, in which my co-authors and I reported the effect of secondary school on HIV status in the Botswana AIDS Impact Surveys II (2004) and III (2008) samples (De Neve et al. 2015). Here, I expand on that analysis in two ways: first, in addition to cohort variation in exposure, I exploit differences in access to grade ten education (senior secondary school) during the pre-reform period using data on village of birth from the complete Botswana Censuses of 2001 and 2011 (Central Statistics

Office 2001, 2011) to provide estimates of impacts on childbearing. Second, using the detailed information contained in the AIDS Impact Surveys, I assess a wide range of possible mechanisms including: sexual behaviors, HIV knowledge and attitudes, and labor force participation.

The paper proceeds as follows: Section 1.2 describes the policy reform and study context. Section 1.3 describes the main datasets used in the analysis. Section 1.4 presents two empirical approaches using cohort and cohort by birth village variation. Section 1.5 presents the main results for HIV and childbearing, using the BAIS datasets, along with robustness checks. Section 1.6 presents results for childbearing by using the Census datasets. In Section 1.7, I explore mediating pathways from education to HIV risk. Section 1.8 concludes.

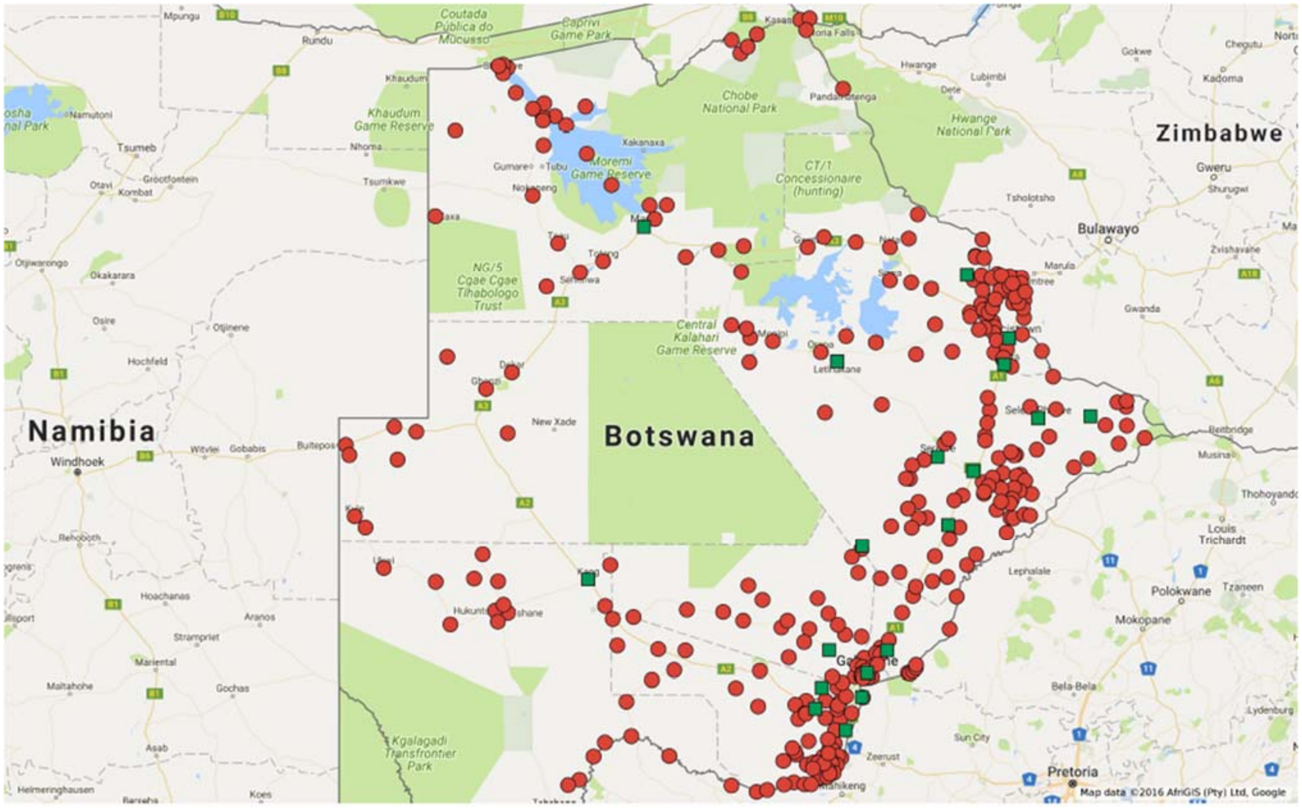
1.2 Policy Reform and Study Context

Botswana is a sparsely-populated, land-locked country in southern Africa, with an economy dependent on diamond-mining. Botswana has among the highest rates of HIV in the world, with 24.6% of adults aged 15-49 years infected in 2013 (BAIS 2013).

Government K-12 education is divided into primary, junior secondary and senior secondary schools. Attendance in primary and junior secondary is compulsory and many students exit after receiving their junior secondary certificate. In 1994, a National Commission on Education (NCE) brought up several concerns with the existing '7+2+3' grade structure, in which two years of junior and three years of senior secondary school followed seven years of primary school. Two years of junior secondary education was deemed insufficient to prepare students for work or further training, and the NCE recommended a switch to a '7+3+2' structure with primary and junior secondary school forming ten years of basic education (rather than nine). Botswana's public education system is strongly centralized and the policy reform was implemented rapidly throughout the country. Fewer than 1% of secondary school students attend private schools (Government of Botswana 2013).

In January 1996, Botswana shifted the tenth year of education from senior secondary to junior secondary school, with the aim to increase access to grade ten (Government of Botswana 1994). The reform may have influenced educational attainment through multiple channels. First, the reform led to a large increase in the supply of grade ten education. There are about nine junior secondary schools for every senior secondary school in Botswana (CSO 2010), which are sparsely located around the country (Figure 1.1); moving grade ten to junior secondary school thus increased the number of seats and reduced travel time for students. Second, it increased the number of years of schooling required to obtain a Junior Certificate, raising the benefits of completing grade ten and establishing continuity with grade nine. Third, if education affects personal preferences for later schooling (e.g., some students may discover that they like school), then increasing grade ten completion could increase progression through grades 11 and 12 in secondary school and even tertiary education.

Fig. 1.1: *Botswana Villages With and Without a Senior Secondary School in 1991 (Pre-Reform)*



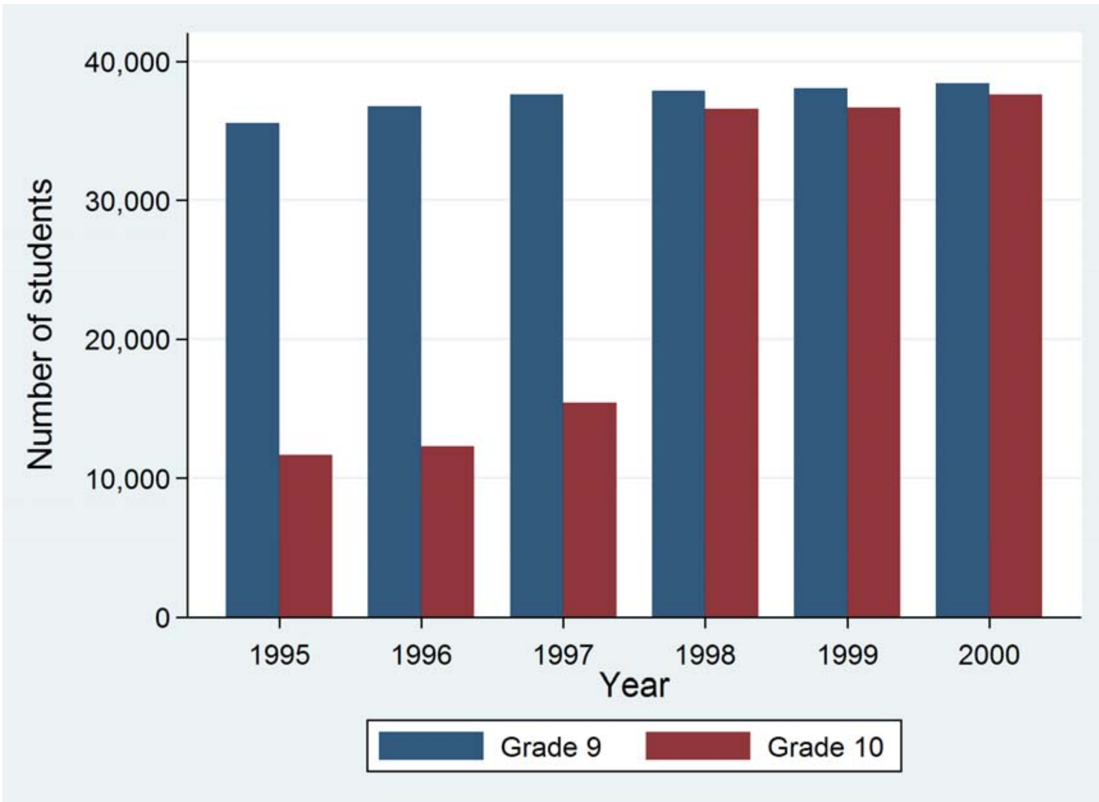
Notes: Figure shows villages in Botswana with (green squares) or without (red circles) a senior secondary school during the pre-reform period. Source: 1991 Population and Housing Census: Guide to the Villages of Botswana, Botswana Central Statistics Office.

The NCE also recommended some curriculum changes in junior and senior secondary school, however the adopted changes were scheduled to be implemented more gradually (over five to six years) and, in some cases, was left to the discretion of the schools themselves (Gaotlhobogwe 2010). It is worth noting, however that at the junior secondary level, which included grade ten post-reform, the NCE plan sought to increase the vocational orientation of academic subjects (a basic computer course) and to increase career guidance to support junior secondary school graduates in their future occupations and training programs. At the senior secondary level, some additional emphasis was placed on practical and professional skills (e.g., computer literacy) and formal links with the private sector (e.g., through internships) (Government of Botswana 1994).

The result of this reform was a very large increase in the number of students attending grade 10 between 1997 and 1998, as reported in Government Education Statistics (Figure 1.2). This was achieved in part through an expansion of capacity, with additional teachers (Figure 1.3a) and classes (Figure 1.3b), to handle the influx of students. A previous study has shown that the policy reform led to an increase in years of schooling – particularly in grade ten completion – across birth cohorts (Borkum 2009).

This 'natural experiment' provides an opportunity to estimate the causal effect of secondary education on HIV risk, childbearing, and possible mediators by comparing outcomes across birth cohorts with differential exposures to the reform.

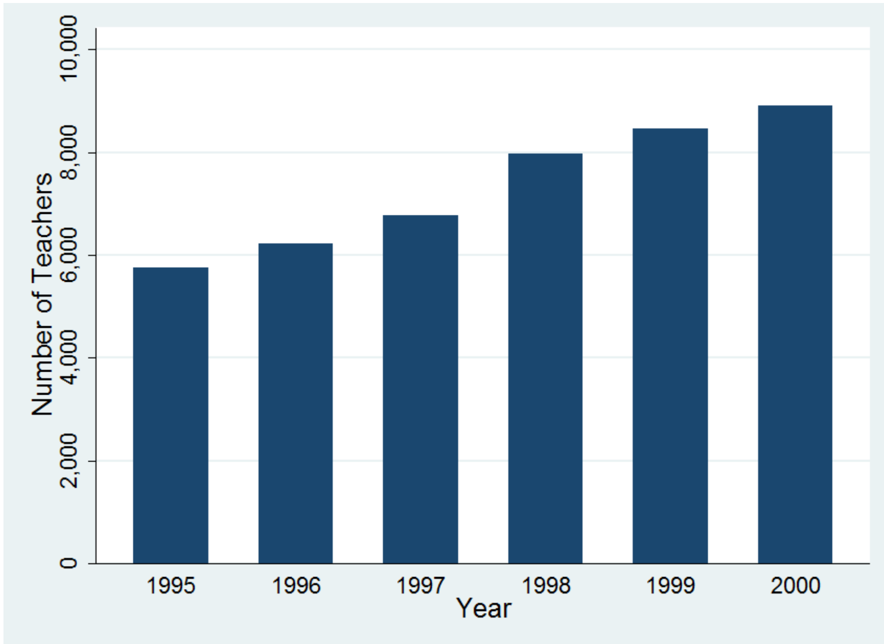
Fig. 1.2: *Enrollment in Ninth and Tenth Grade in Botswana, 1995-2000*



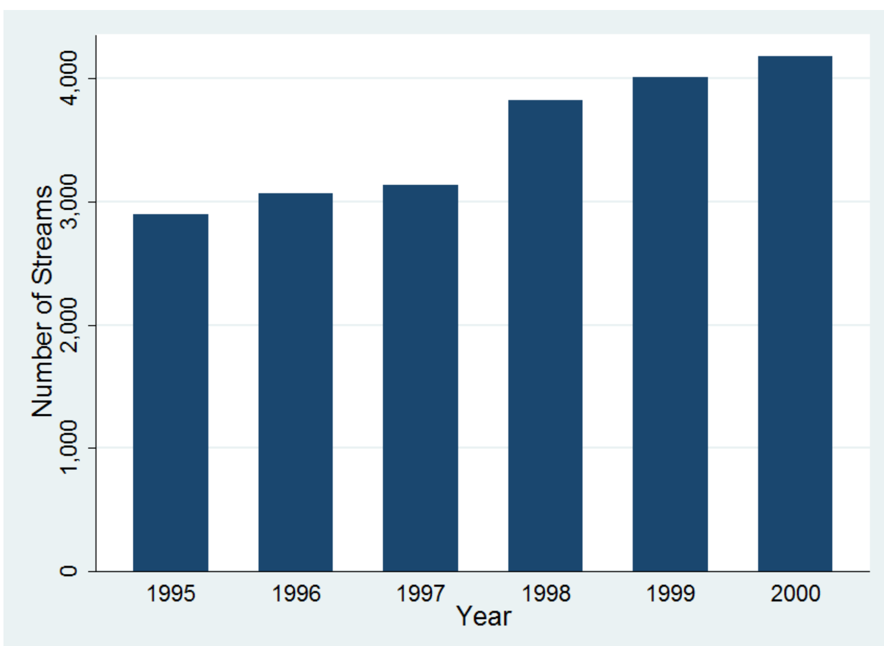
Notes: Number of students in grade nine and ten by year in Botswana between 1995 and 2000. Source: Annual Botswana Education Statistics 1995 - 2000.

Fig. 1.3: *Increased Secondary School Capacity to Handle Student Influx*

(a)



(b)



Notes: Source: Annual Botswana Education Statistics 1995 - 2000.

1.3 Data Description

1.3.1 Study Population

The study population included all women and men living in Botswana, at least 18 years old at the time of the surveys, and who were citizens of Botswana born in or after 1975. Respondents younger than 18 years old at the time of the surveys were excluded because they would not have had the opportunity to complete secondary education. Multiple previous school reforms led to rapid changes in the distribution of schooling for cohorts born before 1975 (Government of Botswana 1994) so they were excluded from the analysis, resulting in a study population ages 18 to 32 years in the BAIS surveys, and a study population ages 18 to 36 in the Census datasets. Immigrants to Botswana were excluded because they would not have been exposed to the education policy reform if they migrated after childhood.

1.3.2 Botswana AIDS Impact Surveys

HIV data were obtained from the Botswana AIDS Impact Surveys (BAIS) II (2004) and III (2008), nationally representative, cross-sectional, population-based household surveys with HIV biomarker collection. BAIS II and III each employed a two-stage probability sample design stratified according to district and major urban centers. For each survey, a representative probability sample of approximately 8,300 households was selected. All household members aged 10-64 who were usual members of the household, and spent the last night in the household, were eligible to be interviewed. For survey year 2004, 8,206 female and 6,656 male respondents were interviewed; in 2008, 7,497 female and 6,055 male respondents were interviewed. Data on demographics, HIV biomarkers, and self-reported sexual behaviors were collected for all respondents. Detailed sampling plans and HIV testing procedures are available from survey final reports (Central Statistics Office 2004, 2008b). Household and individual participation rates were, respectively, 92% and 93% for survey year 2004, and 87% and 82% for survey year 2008. HIV test participation rates were 61% for survey year 2004, and 67% for survey year 2008. A small number of HIV tests (< 2%) produced invalid results. Data on years of schooling completed were available for 99.7% of respondents with an HIV test result.

Data on educational attainment, age in years, gender, citizenship, district of birth, HIV status, childbearing, marriage, sexual and HIV testing behaviors, HIV knowledge and attitudes, employment, and literacy were extracted from the BAIS datasets. Year of birth was calculated based on age and survey year. Primary endpoints were (1) HIV status, which reflects the cumulative probability that a respondent acquired HIV up to his or her age at the time of the BAIS surveys, and (2) childbearing outcomes, including the probability of ever given birth and the number of births at the time of the survey. As secondary endpoints, I also assessed the causal effect of secondary schooling on a range of potential mediating pathways including: age at first intercourse, an important predictor of HIV risk (Pettifor et al. 2004), age at first marriage, years of premarital sex, current sexual and HIV testing behaviors (such as age and number of partners), HIV knowledge and attitudes, literacy, and labor force participation. The text of the specific questions is provided in the Appendix.

1.3.3 Botswana Population Censuses

Data on age in years, gender, birthplace, citizenship, years of schooling completed were obtained from the 2001 and 2011 Censuses, complete censuses of the population of Botswana (Central Statistics Office 2001, 2011). Year of birth was calculated based on age and survey year. There were 1,680,863 persons and 2,038,228 persons enumerated in Botswana during the 2001 and 2011 Census, respectively. Data on age, gender, district of birth, and educational attainment was available for 90.1% of Botswana citizens, yielding a total of 1,512,903 respondents in the 2001 Census and a total of 1,801,450 respondents in the 2011 Census.

1.3.4 Botswana Labor Force Surveys

I extracted data on age, gender, birthplace, citizenship, years of schooling completed, monthly wages and benefits, and hours worked from the Botswana Labor Force Surveys 1995/96 and 2005/06, nationally representative samples with data on economic activity. Year of birth was calculated based on age and survey year. Sampling plans and statistical procedures are available in survey final reports (Central Statistics Office 2008a).

1.4 Empirical Approach

1.4.1 Empirical Approach 1: Variation Across Birth Cohorts

As my first approach, I exploited differential access to grade ten education by year of birth. Historically in Botswana, children were eligible to begin primary school in the year of their seventh birthday. Presuming they would progress on time through school, without repeating or skipping grades, they would enter 10^{th} grade in the year of their 16^{th} birthday. I defined exposure to the reform ("reform cohort") as one if the respondent was aged 16 years or younger in 1996 (born in or after 1981), and zero otherwise.

First, I estimated the effect of exposure to the reform on total years of schooling completed in multivariate OLS regression models ("first stage") and on the distribution of years completed. Second, I assessed the "intention-to-treat" (ITT) effect of being in a reform cohort on my outcomes of interest. Third, I estimated the effect of schooling on outcomes in two-stage least squares (2SLS) regression models, using exposure to the reform as an instrument. I estimated models of the form:

$$(1) \qquad Prob(YearsofSchooling_i \mid X_i, ReformIndicator_i) = \alpha\ ReformIndicator_i + \boldsymbol{\beta}\ \boldsymbol{X_i}$$

$$(2) \qquad Prob(HIVpos_i{=}1 \mid X_i, YearsofSchooling\_hat_i) = \gamma\ YearsofSchooling\_hat_i + \boldsymbol{\delta}\ \boldsymbol{X_i}$$

where $\boldsymbol{X_i} = [Age_i, f(YearofBirth), DistrictofBirth_i]$

In this model, γ estimates the causal effect of schooling on HIV-infection risk if the error terms in the linear models (1) and (2) are independent. In all models, I controlled flexibly for age with a full set of single-year age indicators, to account for the lower expected years of schooling for individuals at younger ages and the non-monotonic pattern of HIV infection across ages in Botswana (Government of Botswana 2009). I included indicators for district of birth. Finally, I adjusted for a continuous linear term in year of birth, to account for continuous trends in educational attainment, HIV infection risk, and childbearing across birth cohorts. When pooling sexes, I added indicators for sex and the interactions of sex with the other covariates**.** Exposure to the policy reform was modeled as an intercept shift for cohorts born in or after 1981. All IV estimates are interpreted as "local" to the subpopulation who "complied" with their treatment assignment – e.g., persons who increased their years of schooling because of the reform (Imbens and Angrist 1994). I used this cohort approach to assess the effect of the reform on outcomes in the BAIS surveys as well as in the Census datasets.

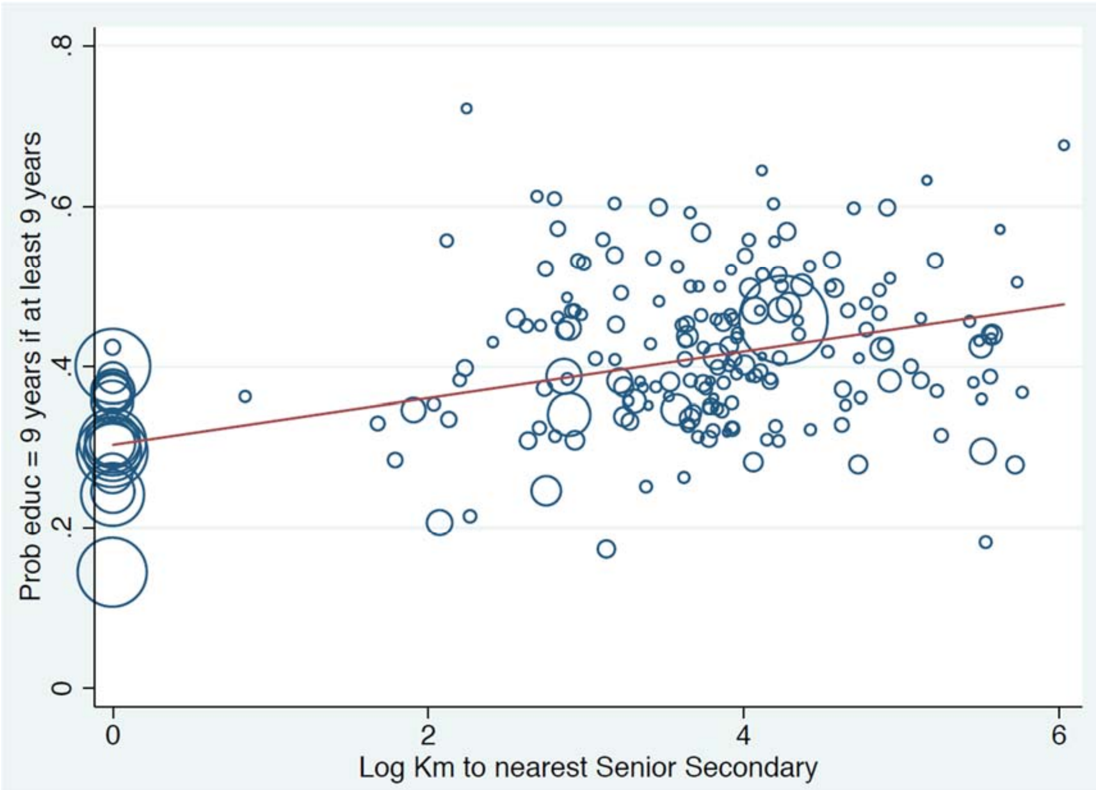
I subjected my main results to a wide range of robustness checks to assess sensitivity to functional form of f(YOB), non-response in the HIV surveys, and the possibility of other contemporaneous changes. These are described below.

1.4.2 Empirical Approach 2: Variation Across Cohorts by Birth Village

As an alternative identification strategy, I exploited differential impact of the reform by birth village using data from the Botswana Census of 2001 and 2011. In the BAIS data, I was only able to impute district of birth for respondents who had never moved. District and village of birth, however, were available in the Census data. Persons who would have dropped out after nine years of schooling in the absence of the policy change are the most likely to have been affected by the reform. Thus, the reform would be expected to have the biggest impact in birth villages where a higher proportion of students completed exactly nine years of schooling pre-reform. I calculated the probability of completing exactly nine years of schooling for each birth village for the pre-reform cohorts born 1975-1978, Pr(Educ=9), excluding 1979 and 1980 cohorts from these estimates because of potential exposure to the reform among students who entered school late or were kept back.

Observed village estimates of Pr(Educ=9) may contain idiosyncratic errors which may be subject to mean reversion in the post-reform period. To eliminate potential bias that may result, I used (approximately) time-invariant supply-side variables in the pre-reform period to instrument for grade nine and sorted villages by predicted probabilities of exactly nine years of schooling pre-reform, Pr(Educ=9)_hat. I use three supply-side instruments: first, I assessed the presence of a senior secondary school in a respondent's village of birth in the pre-reform period and calculated the log(distance (km)) to the nearest senior secondary school for each village of birth. Figure 1.1 displays a map of villages in the pre-reform period in Botswana, with and without a senior secondary school. Figure 4 displays the relationship between distance to a secondary school and the proportion dropping out after grade nine among those reaching grade nine. Second, I calculated the ratio of the number of available secondary school classrooms divided by the age-eligible population for each district of birth in the pre-reform period. Third, I calculated the ratio of the number of secondary school classes divided by the age-eligible population for each district of birth in the pre-reform period. I predicted the probability of completing exactly nine years of schooling by village of birth in the pre-reform period based on my three structural variables, while controlling for age fixed effects, using the data from the Botswana Census of 2001. This measure of susceptibility to the reform, interacted with being born in a post-reform cohort, provides my source of identifying variation.

Fig. 1.4: *Grade 9 Completion and Distance to Senior Secondary School*



Notes: Figure shows the naïve association between the probability of completing exactly nine years of schooling and log(distance to nearest senior secondary school) pre-reform, conditional on completing at least nine years of schooling. Source: Botswana Census 2011.

I estimate two different types of models under Empirical Approach 2. First, using only the 2011 Census, I estimated difference-in-differences 2SLS models, comparing people who were older and younger (different birth cohorts) across birth villages with different susceptibility to the reform, as measured by Pr(Educ=9)_hat. I controlled flexibly for single-year age indicators (equivalent to birth cohort indicators in this single Census wave) and birth village fixed effects. I estimated the following:

$$(1) \quad Prob(YearsofSchooling_i \mid X_i, ReformIndicator*BirthvillageEduc9\_hat_i) = \alpha\ ReformIndicator*BirthvillageEduc9\_hat_i + \beta_i + \gamma_i$$

$$(2) \quad Prob(HIVpos_i{=}1 \mid X_i, ReformIndicator*BirthvillageEduc9\_hat_i) = \delta\ YearsofSchooling\_hat_i + \eta_i + \theta_i$$

where β_i and η_i are age fixed effects and γ_i and θ_i are birth village fixed effects.

Second, the availability of two Census years also allowed me to disentangle age and cohort effects. Using data from both the 2001 and 2011 Census, I estimate models with a triple-differences flavor, in which I control for single-year age fixed effects, village of birth fixed effects, a post-reform indicator, and also adjust for a continuous term in year of birth (noting that birth cohort = survey_year – age). Identification comes from the interaction of the reform indicator with Pr(Educ=9)_hat measured at the birth village level. When including both Census waves, I estimated models of the form:

$$(1) \quad Prob(YearsofSchooling_i \mid X_i, ReformIndicator*BirthvillageEduc9\_hat_i) = \alpha\ ReformIndicator*BirthvillageEduc9\_hat_i + \pi\ YearofBirth_i + \beta_i + \mu_i + \gamma_i$$

(2) $Prob(HIVpos_i=1 | X_i, ReformIndicator*BirthvillageEduc9\_hat_i) =$
$\delta\ YearsofSchooling\_hat_i + \sigma\ YearofBirth_i + \eta_i + \Omega_i + \theta_i$

where β_i and η_i are age fixed effects, μ_i and Ω_i is a post-reform indicator, and γ_i and θ_i are village of birth fixed effects, and the difference compared to the first approach is the inclusion of a linear term in $YearofBirth_i$ and post-reform indicator.

To assess the robustness of my results, I conducted a range of sensitivity analyses. In addition to running each model separately using either the Census 2011 or using both Census years, I also excluded respondents born in Gaborone—the capital of Botswana and its largest city. I included district-specific time trends using the interaction of district of birth with a linear trend in year of birth. Finally, I included the interaction of the reform indicator with district of birth fixed effects to identify off of variation within districts only. In all models, standard errors were clustered at the village of birth level.

Stata (version 14.0, StataCorp, College Station, Texas) was used for all statistical analyses. This study was reviewed by the Harvard T.H. Chan School of Public Health Institutional Review Board and considered exempt from full review as it was based on anonymous data with no identifiable information on survey participants.

1.5 BAIS Results: HIV Risk, Childbearing

1.5.1 Sample Description BAIS Data

The 2004 and 2008 BAIS surveys included 3,965 women and 3,053 men with valid HIV biomarkers, for a total of 7,018 respondents. Table 1.1 shows summary statistics. Mean age was 22.7 (SD 3.1) for women and 22.6 (SD 3.2) for men in the BAIS II survey and 24.9 (SD 4.2) for women and 24.7 (SD 4.3) for men in the BAIS III survey. Mean years of schooling was similar for men and women, at about ten years. Age at first intercourse was 18.0 (SD 2.0) years for women and 17.8 (SD 2.5) years for men in the BAIS II survey, and 18.2 (SD 2.5) for women and 18.5 (SD 3.0) for men in the BAIS III survey. 28% of women and 11% of men in the BAIS II study sample and 27% of women and 12% of men in the BAIS III study sample were HIV positive.

Table 1.1: *Summary Statistics for the BAIS and Census Datasets*

Survey (Year)	BAIS II (2004)		BAIS III (2008)		Census (2001)		Census (2011)	
Subsample	Female	Male	Female	Male	Female	Male	Female	Male
Age, Mean (SD)	22.7 (3.1)	22.6 (3.2)	24.9 (4.2)	24.7 (4.3)	21.8 (2.6)	21.7 (2.6)	26.4 (5.3)	26.4 (5.4)
Years of Schooling, Mean (SD)	10.0 (3.0)	9.7 (4.0)	10.5 (3.2)	10.3 (3.8)	9.5 (3.1)	9.0 (3.7)	10.6 (3.3)	10.2 (3.7)
At Least 10 Years of Schooling, %	62.4	65.2	72.6	73.0	55.4	57.1	74.8	73.5
HIV Positive, %	28.3	11.1	27.3	12.4				
Ever Had Sex, %	88.2	77.9	92.7	83.1				
Age at First Intercourse,Mean (SD)	18.0 (2.0)	17.8 (2.5)	18.2 (2.5)	18.5 (3.0)				
Ever Given Birth, %	60.6		72.1		53.6		64.9	
Number of Children, Mean (SD)	0.99 (1.03)		1.39 (1.42)		0.83 (0.98)		1.36 (1.44)	
Ever Married, %	4.9	1.0	7.1	2.6	4.0	1.1	10.2	5.4
Literacy, %	83.0	80.0	91.1	86.0				
Labor Force Participation, %	59.9	72.5	62.8	75.0	51.8	62.3	64.7	76.1
Birth Villages, No					464	464	464	464
Observations	1,760	1,354	2,205	1,699	159,529	143,901	342,884	322,613

Notes: The Census samples include survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and born in or after 1975. The BAIS sample includes citizens of Botswana, at least 18 years old, born in or after 1975, and with a valid HIV test result. Total N with Age at First Intercourse was 1,520 for women and 1,012 for men in BAIS II (2004), and 1,987 for women and 1,348 for men in BAIS III (2008). Sample weights used as provided in the BAIS samples. Source: Botswana AIDS Impact Survey II (2004) (N: 15,479) and III (2008) (N: 14,127), and Botswana Census 2001 (N: 1,512,903) and 2011 (N: 1,801,450).

1.5.2 Descriptive Association between Education and HIV Infection Risk

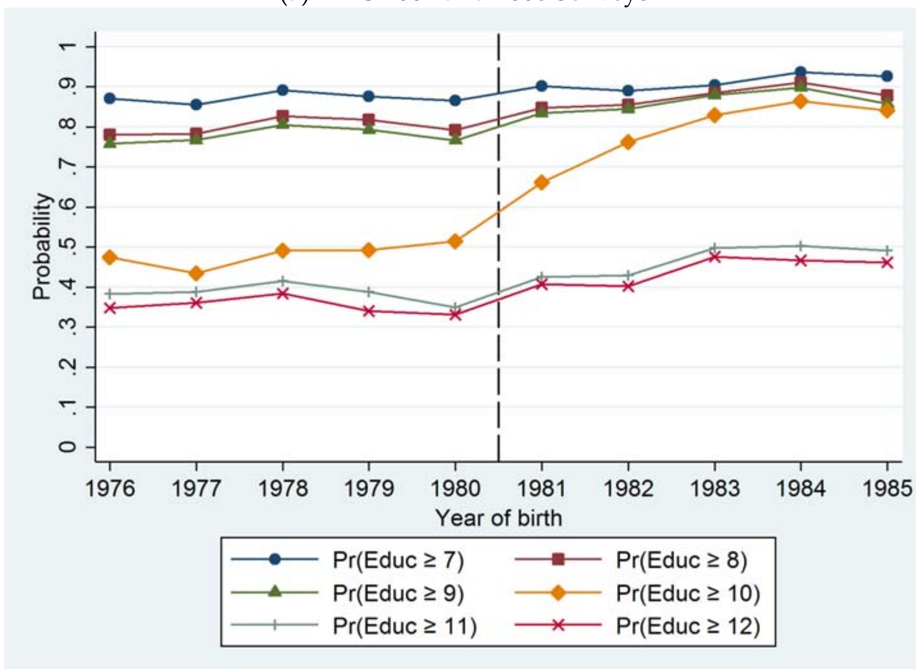
Figure A1 in the Appendix shows the crude associations between education and HIV infection risk for men and women. The relationship between schooling and HIV infection risk appears non-monotonic. HIV risk peaks for persons completing 8 - 9 years of education, and declines sharply after nine years of schooling. Table A1 in the Appendix shows the covariate-adjusted OLS association between years of schooling and HIV infection risk. On average, one additional year of schooling is associated with 1 - 2 percentage points lower HIV risk. However, each additional year of schooling up to nine years was associated with a 0.3 percentage point *higher* risk for HIV infection (p = 0.229); by contrast, each additional year of schooling above nine years of schooling was associated with a -3.6 percentage point lower risk of HIV infection ($p < 0.001$). Since these associations may be confounded by unobserved characteristics, I used an instrumental variables approach to obtain causal effect estimates.

1.5.3 Effect of the 1996 Grade Reform on Years of Schooling Completed

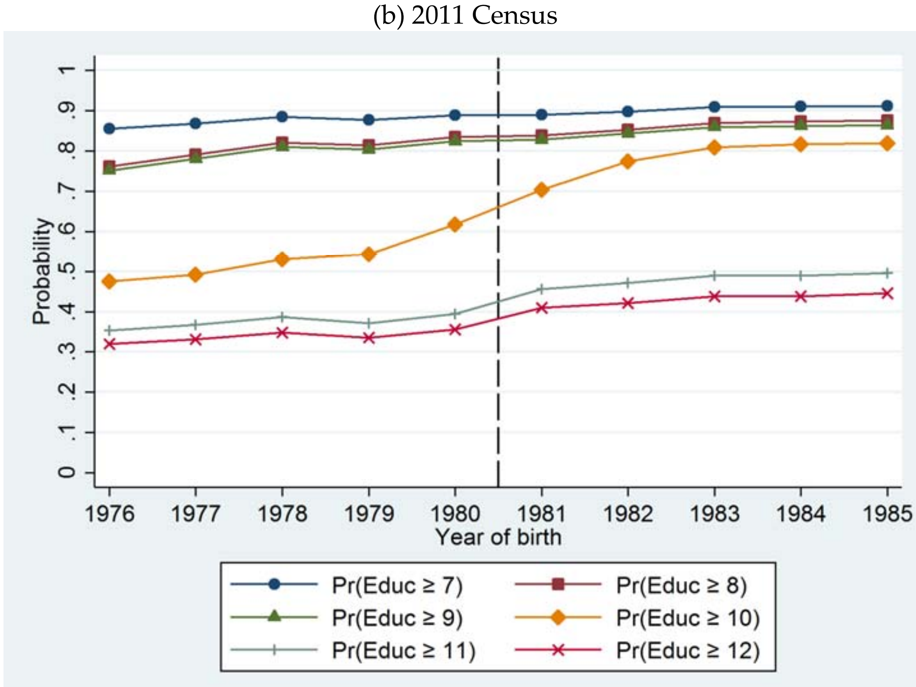
Figure 1.5 shows the proportion of respondents who completed at least 7, 8, 9, 10, 11, or 12 years of schooling and how this distribution changed across birth cohorts. The proportion of students completing at least 7, 8, or 9 years of schooling rose gradually and continuously across birth cohorts. However, the proportion of students with at least ten years of schooling increased sharply for those birth cohorts affected by the reform. Modest increases in completion of 11 and 12 years of schooling were also observed for the reform cohorts. Tables 1.2 and A2 in the Appendix display regression estimates of the impact of the reform on educational attainment. In the BAIS sample, the reform induced an increase of 0.79 years of schooling ($p < 0.001$); 0.64 for women (p = 0.004) and 1.01 for men (p = 0.002).

Fig. 1.5: *Educational Attainment by Birth Cohort in Botswana*

(a) BAIS 2004 and 2008 Surveys



(b) 2011 Census



Notes: Pr(Educ $\geq$ X) is the probability that the respondent has attained at least X years of schooling. Sample includes survey respondents who were citizens of Botswana, at least 22 years old at the time of the surveys and had a valid HIV test result. Survey weights used as provided. Individuals born in 1981 or later (dotted line) were classified as exposed. Source: Botswana AIDS Impact Survey II (2004) and III (2008) and Botswana Census 2011.

1.5.4 Causal Effect of Education on HIV Infection Risk and Childbearing

Table 1.2 presents ITT results, regressing HIV status directly on the instrument (cohort exposure to the reform) and covariates. Women who were exposed to the policy reform were 7.4 percentage points less likely to be HIV positive (p = 0.017); men were 5.0 percentage points less likely to be HIV positive (p = 0.052). In the pooled sample, the coefficient was 6.4 percentage points (p = 0.002). Table 1.2 also shows 2SLS results for the effect of additional years of secondary schooling on HIV status. These instrumental variable estimates show that additional years of schooling induced by the reform had a protective effect against HIV infection. Batswana who stayed in school for an additional year had an 8.1 percentage point lower risk of HIV infection (p = 0.008); 11.6 percentage points for women (p = 0.045) and 5.0 percentage points for men (p = 0.085). Although effects were generally larger in absolute terms for women than for men, I could not reject the null hypothesis that the effect of the policy change was the same for men and women (coefficient on interaction = 0.023, p = 0.556). Education also reduced the proportion of women that had ever given birth by 16.0 percentage points (p = 0.006). Since the reform increased completion of grades ten and above, I interpret the instrumental variable estimates as the causal effect of additional secondary school education on HIV risk and childbearing, for the population that would have otherwise dropped out of school after grade nine.

Table 1.2: *Empirical Strategy 1 (BAIS 2004 and 2008): HIV Infection Risk and Childbearing*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Model	**First Stage**	**ITT**	**ITT**	**ITT**	**2SLS**	**2SLS**	**2SLS**
Dependent Variable	**Schooling (years)**	**HIV-Positive**	**Ever Given Birth**	**Number of Children**	**HIV-Positive**	**Ever Given Birth**	**Number of Children**
				Female Subsample			
Coefficient on Reform Indicator (Models 1 - 4)	0.635***	-0.074**	-0.114***	-0.181**	-0.116**	-0.158***	-0.246**
or Years of Schooling (Models 5 - 7)	(0.223)	(0.031)	(0.033)	(0.079)	(0.058)	(0.057)	(0.107)
Observations	3,965	3,965	3,644	3,656	3,965	3,644	3,656
R-squared	0.034	0.095	0.198	0.228			
Probability Dependent Variable, Pre-Reform		0.323	0.785	1.38			
				Male Subsample			
Coefficient on Reform Indicator (Models 1 - 2)	1.005***	-0.050*			-0.050*		
or Years of Schooling (Model 5)	(0.322)	(0.026)			(0.029)		
Observations	3,053	3,053			3,053		
R-squared	0.033	0.070					
Probability Dependent Variable, Pre-Reform		0.168			0.164		
				Both Sexes			
Coefficient on Reform Indicator (Models 1 - 2)	0.792***	-0.064***			-0.081***		
or Years of Schooling (Model 5)	(0.188)	(0.021)			(0.031)		
Observations	7,018	7,018			7,018		
R-squared	0.036	0.123					
Probability Dependent Variable, Pre-Reform		0.255					

Notes: Regressions 1 to 4 are OLS models; regressions 2 - 4 are Intent-to-Treat (ITT) models. Regressions 5 - 7 are two-stage-least-squares (2SLS) models in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave, and indicators for district of birth. Regressions for the subsample with both sexes additionally control for sex and the interactions of sex with each covariate. The instrument was defined as = 1 if year of birth > 1980. F-statistics were between 8.6 and 10.4 (models 5 - 7). Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Sample includes citizens of Botswana, at least 18 years old, born in or after 1975, and with a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

1.5.5 Robustness Checks

In general, results were robust to a range of sensitivity analyses (see Table A5 in the Appendix for results in the pooled sample and sex-specific results). I assessed the robustness of my results to the presence of non-linearities in long-run cohort trends in HIV risk, by controlling for a quadratic in year of birth in addition to the linear term included in the main analysis, and find almost identical results (Table A5, columns 4, 5). The assumption that underlying trends are approximately linear is also more plausible the narrower the window of cohorts included. Using narrower window of birth cohorts – 1981 +/- 4, 5 years, results remain similar, although the estimates are less precise (Table A5, columns 7, 8). HIV consent rates were incomplete and it is possible that selective non-response may have biased my estimates. However, exposure to the reform was not significantly associated with consent rates. Further, imputing for HIV status based on observables, my results were essentially unchanged (Table A5, columns 10, 11). To rule out the possibility that other national policy changes might have affected HIV risk for the same cohorts, I used an alternative identification strategy exploiting the fact that the reform would be expected to have the biggest impact in districts where a higher proportion of students completed exactly nine years of schooling in the pre-reform period. I created an indicator for high vs. low proportion of 9^{th} grade completion and interacted this variable with a reform indicator to implement a difference-in-differences strategy and found similar results (Table A5, column 12). Finally, I conducted a placebo test. Dividing the sample into respondents with at least 9 years of schooling and respondents with less than 9 years of schooling, I estimated the effect of the policy reform among those with at least 9 years – i.e. those people whose schooling increased due to the reform; and those with less than nine years of schooling, who experienced no increase in years of schooling (Table A5, columns 14, 15). Further details and additional robustness checks are provided in the Appendix. Table A5 in the

Appendix displays robustness checks in the pooled sample, and for men and women separately.

1.6 Census Results: Childbearing

1.6.1 Sample Description Census Data

The 2001 and 2011 Census included 502,413 women and 466,514 men satisfying my inclusion criteria (YOB ≥ 1975, age ≥ 18) for a total of 968,927 respondents. Table 1.1 shows summary statistics. Mean age was 21.8 (SD 2.6) for women and 21.7 (SD 2.6) for men in the 2001 Census and 26.4 (SD 5.3) for women and 26.4 (SD 5.4) for men in the 2011 Census. Mean years of schooling was about ten years. 54% of women in the Census 2001 study sample and 65% of women in the Census 2011 study sample had given birth.

1.6.2 Cohort Effects

Figure 1.5 and Table 1.3 shows first stage, ITT, and 2SLS results for childbearing using the Census data. Similar to results using the BAIS datasets, the reform had a large effect on educational attainment (in the pooled sample: 0.346 additional years, $p < 0.001$). Among women, exposure to the reform was associated with a decrease of 2.4 percentage points ($p < 0.001$) in the probability of ever given birth, and with having 0.094 less children ($p < 0.001$) by the time of the Census. Each additional year of secondary schooling reduced the probability of ever given birth by 5.7 percentage points ($p < 0.001$) and reduced the number of children by 0.228 ($p < 0.001$). These results are consistent with the instrumental variable estimates using the BAIS datasets.

Table 1.3: *Empirical Strategy 1 (Census 2001 and 2011): Childbearing*

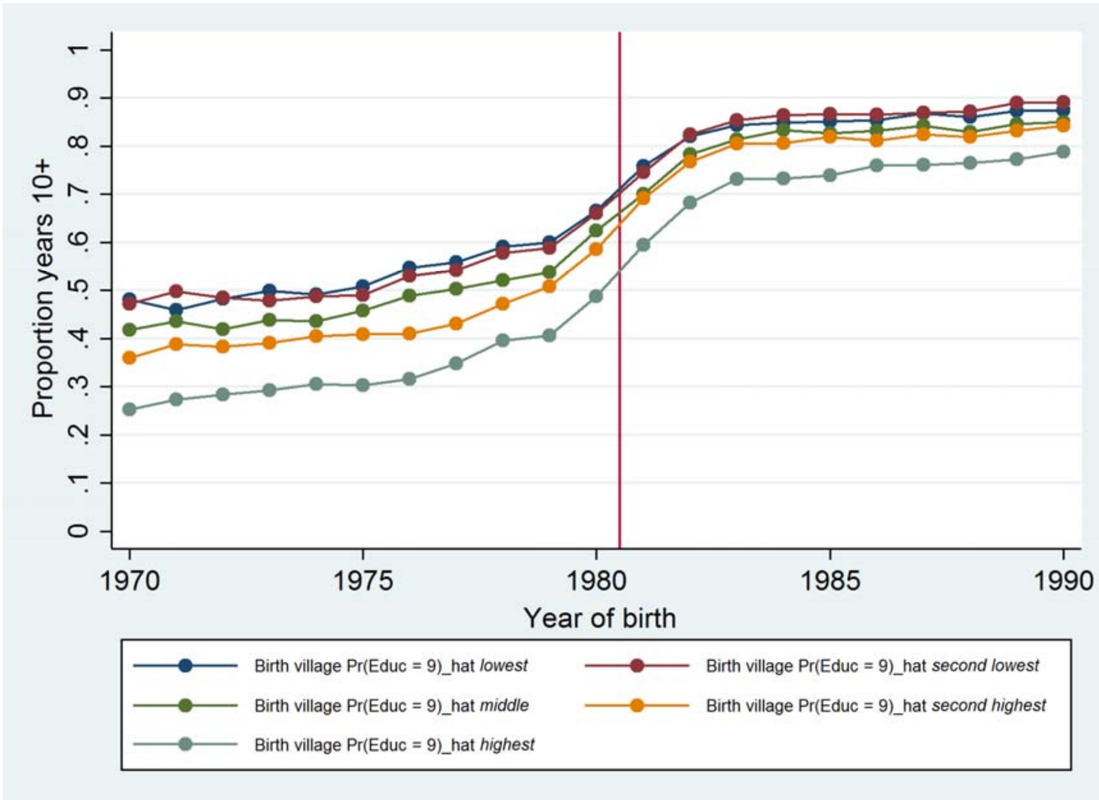
	(1)	(2)	(3)	(4)	(5)
Model	**First Stage**	**ITT**	**ITT**	**2SLS**	**2SLS**
Dependent Variable	**Schooling (years)**	**Ever Given Birth**	**Number of Children**	**Ever Given Birth**	**Number of Children**
			Female Subsample		
Coefficient on Reform Indicator (Models 1 - 3)	0.410***	-0.024***	-0.094***	-0.057***	-0.228***
or Years of Schooling (Models 4 - 5)	(0.022)	(0.003)	(0.005)	(0.007)	(0.014)
Observations	500,123	500,123	500,123	500,123	500,123
R-squared	0.103	0.259	0.310		
Probability Dependent Variable, Pre-Reform		0.676	1.33		

Notes: Regressions 1 to 3 are OLS models, of which regressions 2 - 3 are Intent-to-Treat (ITT) models. Regressions 4 - 5 are two-stage-least-squares (2SLS) models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for census year, and indicators for district of birth. The instrument was defined as = 1 if year of birth > 1980. F-statistics were above 400. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes citizens of Botswana, 18+ years old at the time of the census, and born in or after 1975. Source: Botswana Census 2001 and 2011.

1.6.3 Exploiting Differential Impact by Birth Village

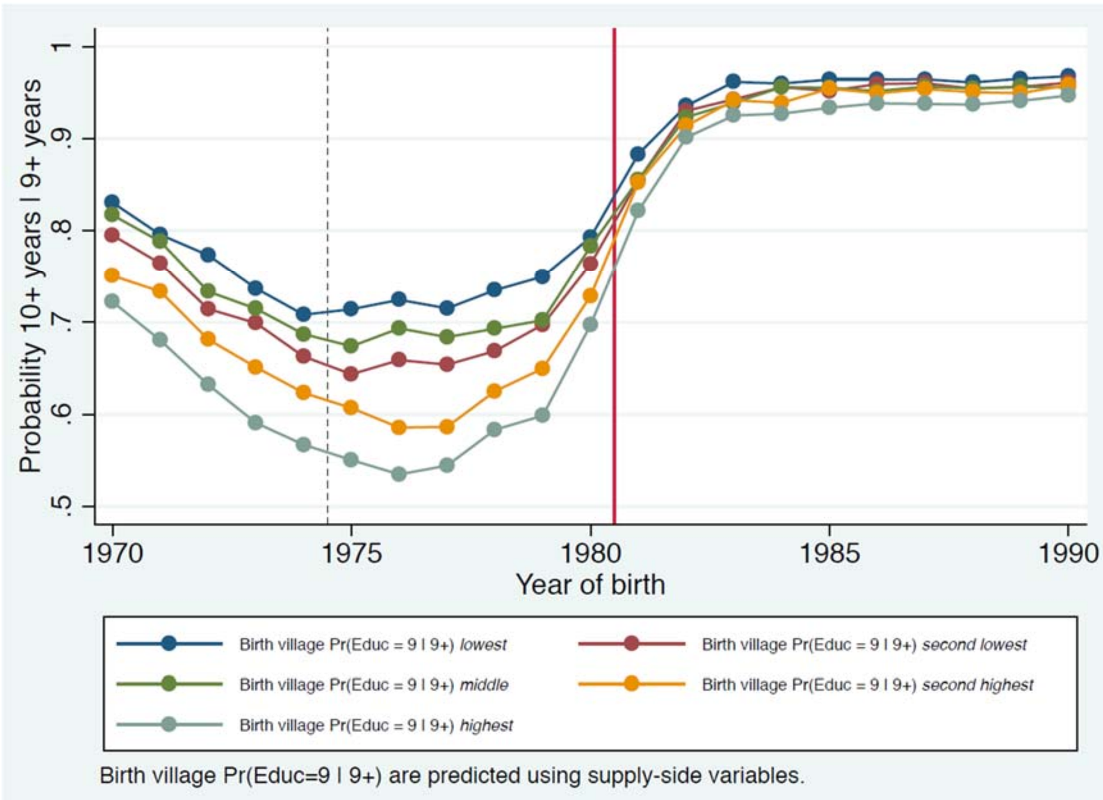
Figure 1.6 displays the probability of completing at least ten years of schooling by year of birth, separately for each birth village quintile of Pr(Educ=9)_hat (the village level proportion of ninth graders progressing to tenth grade in the 1975-1978 birth cohorts, instrumented by the supply-side variables described above). Although the probability of progressing to grade ten increases in all village quintiles, the gap between quintiles shrinks, with larger absolute gains in those villages with the highest proportion of ninth grade dropouts. The source of variation is even more visible when conditioning on reaching nine years of schooling. Figure 1.7 shows completion of at least grade ten conditional on completing grade nine, stratified by quintiles of Pr(Educ=9|Educ≥9)_hat. The large variation in progression to grade ten in the pre-reform period nearly disappears in the post-reform period. Finally, Figure 1.8 displays the difference in completion of at least grade ten years of schooling between the post-reform (YOB 1982 – 1985) and pre-reform (YOB 1975 – 1978) cohorts for each of the birth village Pr(Educ=9)_hat quintiles. The difference in grade ten completion pre- vs. post-reform increases monotonically as the probability of completing exactly nine years of schooling pre-reform increases. Birth villages with a higher predicted proportion of respondents with nine years of schooling were more likely to be affected.

Fig. 1.6: *Grade 10 Completion by Cohort and Predicted Birth Village Grade Nine Completion*



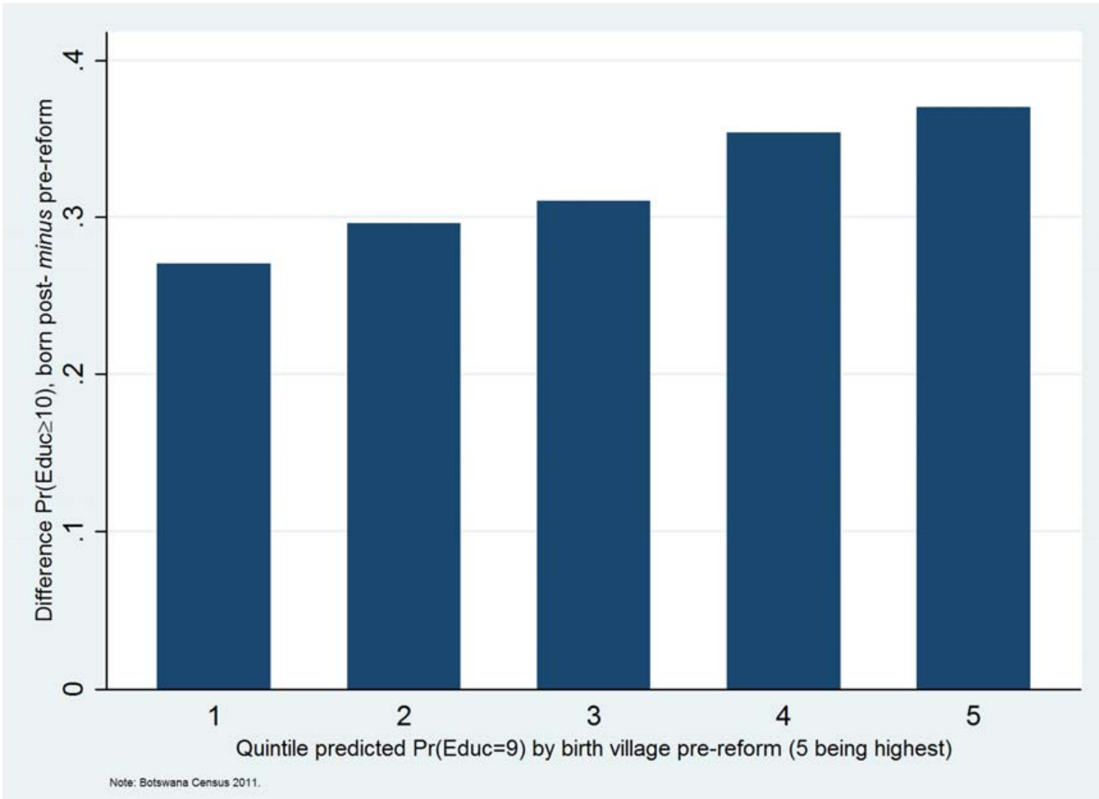
Notes: Figure shows the difference in probability of completing at least ten years of schooling by birth cohort for each Pr(Educ=9)_hat quintile. Pr(Educ=9)_hat was defined as the predicted percentage of ninth grade completion by village of birth for the pre-reform cohorts (YOB 1975 – 1978) based on birth village distance to a senior secondary school, the ratio of classrooms divided by the age eligible population by birth district, and the ratio of class streams divided by the age eligible population by birth district. Source: Botswana Census 2011.

Fig. 1.7: *Grade 10 Completion by Cohort, Conditional on Grade 9 Completion*



Notes: Figure shows the difference in probability of completing at least ten years of schooling by birth cohort for each Pr(Educ=9)_hat quintile. Pr(Educ=9)_hat was defined as the predicted percentage of ninth grade completion by village of birth for the pre-reform cohorts (YOB 1975 – 1978) based on birth village distance to a senior secondary school, the ratio of classrooms divided by the age eligible population by birth district, and the ratio of class streams divided by the age eligible population by birth district. Source: Botswana Census 2011.

Fig. 1.8: *Difference in Grade 10 Completion by Reform Exposure and Birth Village Quintiles*



Notes: Figure shows the difference in probability of completing at least ten years of schooling between the post-reform birth cohorts (YOB 1982 – 85) and pre-reform cohorts (YOB 1975 - 78) for Pr(Educ=9)_hat quintile during the pre-reform period. Pr(Educ=9)_hat was defined as the predicted percentage of ninth grade completion by village of birth for the pre-reform cohorts (YOB 1975 – 1978) based on birth village distance to a senior secondary school, the ratio of classrooms divided by the age eligible population by birth district, and the ratio of class streams divided by the age eligible population by birth district. Source: Botswana Census 2011.

Tables 1.4 – 1.5 show first stage, ITT, and 2SLS results when exploiting differential impact by village of birth. Consistent with empirical approach 1 using the BAIS surveys, additional secondary education reduced childbearing. Each additional year of schooling reduced the number of children by 0.273 ($p < 0.001$) and the probability of ever given birth by 5.5 percentage points ($p < 0.001$).

Table 1.4: *Empirical Strategy 2 (Census 2001 and 2011): First Stage Results*

Model: OLS	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Schooling (years)	Schooling (years)	Schooling (years)	Schooling (years)	Schooling (years)	Schooling (years)	Schooling (years)	Schooling (years)
Coefficient on endogenous variable								
				Subsample: Women				
i.ReformIndicator#c.BirthvillageEduc9_hat	0.115***	0.094***	0.104***	0.084***	0.059***	0.051***	0.090***	0.066***
	(0.018)	(0.016)	(0.015)	(0.013)	(0.013)	(0.010)	(0.016)	(0.013)
Basic Covariates								
i.Age	✓	✓	✓	✓	✓	✓	✓	✓
YOB	-	✓	-	✓	-	✓	-	✓
i.Birthvillage	✓	✓	✓	✓	✓	✓	✓	✓
Additional Controls								
i.ReformIndicator	-	✓	-	✓	-	✓	-	✓
i.Birthdistrict#c.YOB	-	-	-	-	✓	✓	-	-
i.ReformIndicator#i.Birthdistrict	-	-	-	-	-	-	✓	✓
Sample								
Census Year 2011	✓	✓	✓	✓	✓	✓	✓	✓
Census Year 2001	-	✓	-	✓	-	✓	-	✓
Excluding Respondents Born in Gaborone	-	-	✓	✓	-	-	-	-
Observations	317,768	468,031	300,159	444,745	317,768	468,031	317,768	468,031
R-squared	0.114	0.142	0.107	0.134	0.116	0.144	0.115	0.143
				Subsample: Men				
i.ReformIndicator#c.BirthvillageEduc9_hat	0.088***	0.074***	0.073***	0.061**	0.050***	0.039***	0.101***	0.064***
	(0.029)	(0.025)	(0.028)	(0.024)	(0.018)	(0.014)	(0.021)	(0.017)
Basic Covariates								
i.Age	✓	✓	✓	✓	✓	✓	✓	✓
YOB	-	✓	-	✓	-	✓	-	✓
i.Birthvillage	✓	✓	✓	✓	✓	✓	✓	✓
Additional Controls								
i.ReformIndicator	-	✓	-	✓	-	✓	-	✓
i.Birthdistrict#c.YOB	-	-	-	-	-	✓	-	-
i.ReformIndicator#i.Birthdistrict	-	-	-	-	-	-	-	✓
Sample								
Census Year 2011	✓	✓	✓	✓	✓	✓	✓	✓
Census Year 2001	-	✓	-	✓	-	✓	-	✓
Excluding Respondents Born in Gaborone	-	-	✓	✓	-	-	-	-
Observations	298,599	433,893	282,028	412,213	298,599	433,893	298,599	433,893
R-squared	0.108	0.133	0.101	0.125	0.111	0.135	0.110	0.134

All regressions are OLS models. The reform indicator was defined as = 1 if YOB > 1980. BirthvillageEduc9_hat was defined as the predicted percentage of ninth grade completion by village of birth among respondents unexposed to the reform (YOB 1975 - 1978) based on three structural variables (BirthvillageLogDistanceSSS, BirthdistrictRatioClassroomAgeEligible and BirthdistrictRatioStreamsAgeEligible), while controlling for age fixed effects and using the Botswana Census 2001 data. BirthvillageLogDistanceSSS was defined as the log(distance in km) to the closest birth village with a senior secondary school using the Guide to the Villages of Botswana 1991. BirthdistrictRatioClassroomAgeEligible was defined as the ratio of the number of secondary school classrooms divided by the population ages 15 - 19 in the pre-reform period using the Botswana Census 1991 and Education Statistics of 1995. BirthdistrictRatioStreamsAgeEligible was defined as the ratio of the number of secondary school class streams divided by the population ages 15 - 19 in the pre-reform period using the Botswana Census 1991 and Education Statistics of 1995. Standard errors were clustered at the birth village level. *** p<0.01, ** p<0.05, * p<0.1. No weights used. Sample includes citizens of Botswana, 18+ years old at the time of the census, and born in or after 1975. Source: Botswana Census 2001 and 2011.

Table 1.5: *Empirical Strategy 2 (Census 2001 and 2011): ITT and 2SLS Results*

Model	(1) ITT	(2) ITT	(3) ITT	(4) 2SLS	(5) 2SLS	(6) 2SLS
Dependent Variable	**Number of Children**	**Number of Children**	**Ever Given Birth**	**Number of Children**	**Number of Children**	**Ever Given Birth**
			Female Subsample			
Coefficient on i.ReformIndicator#c.BirthvillageEduc9_hat	-0.044***	-0.026***	-0.006***	-0.377***	-0.273***	-0.055***
(models 1 - 3) or schooling (models 4 - 6)	(0.007)	(0.004)	(0.002)	(0.053)	(0.037)	(0.016)
Basic Covariates						
i.Age	✓	✓	✓	✓	✓	✓
YOB	-	✓	-	-	✓	-
i.YOB	-	-	-	-	-	-
i.Birthvillage	✓	✓	✓	✓	✓	✓
i.Birthdistrict	-	-	-	-	-	-
Additional Controls						
i.ReformIndicator	-	✓	-	-	✓	-
Sample						
Census Year 2011	✓	✓	-	✓	✓	-
Census Year 2001	-	✓	✓	-	✓	✓
Observations	318,953	468,563	149,610	317,768	467,231	149,463
R-squared	0.314	0.321	0.211	-	-	-

Regressions 1 - 3 are OLS models. Regressions 4 - 6 are 2SLS models using exposure to the reform as an instrument for years of schooling. The instrument was the interaction of a reform indicator and a continous term in BirthvillageEduc9_hat.The reform indicator was defined as = 1 if YOB > 1980. BirthvillageEduc9_hat was defined as the predicted percentage of ninth grade completion by village of birth among respondents unexposed to the reform (YOB 1975 - 1978) based on three structural variables (BirthvillageLogDistanceSSS, BirthdistrictRatioClassroomAgeEligible and BirthdistrictRatioStreamsAgeEligible), while controlling for age fixed effects and using the Botswana Census 2001 data. BirthvillageLogDistanceSSS was defined as the log(distance in km) to the closest birth village with a senior secondary school using the Guide to the Villages of Botswana 1991. BirthdistrictRatioClassroomAgeEligible was defined as the ratio of the number of secondary school classrooms divided by the population ages 15 - 19 in the pre-reform period using the Botswana Census 1991 and Education Statistics of 1995. BirthdistrictRatioStreamsAgeEligible was defined as the ratio of the number of secondary school class streams divided by the population ages 15 - 19 in the pre-reform period using the Botswana Census 1991 and Education Statistics of 1995. Standard errors were clustered at the birth village level. *** p<0.01, ** p<0.05, * p<0.1. No weights used. Sample includes citizens of Botswana, 18+ years old, and born in or after 1975. Source: Botswana Census 2001 and 2011.

1.7 Causal Pathways from Education to HIV Risk and Childbearing

The BAIS surveys provide detailed information on knowledge, attitudes and behaviors that provide insight into mediating pathways. To investigate the causal pathways from education to HIV risk and childbearing, I assessed the effect of the reform on potential mediators, my secondary endpoints. Tables 1.6 – 1.8 show 2SLS results for the effect of additional years of schooling on age at first intercourse, marriage, sexual and HIV testing behaviors, HIV knowledge and attitudes, literacy, and labor market outcomes for women and men separately. Schooling had no effect on HIV knowledge; however it influenced other HIV risk factors, HIV attitudes, literacy, and labor market outcomes. For both men and women, education increased condom use, HIV testing, and the proportion reporting that is acceptable for women to carry condoms. For women, education delayed sexual debut by 0.76 years (p = 0.004), and increased labor force participation by 17.2 percentage points (p = 0.025). Interestingly, the delay in sexual debut was evident not just at age 16, but all the way to age 20 (Table 1.8), suggesting a change in preferences, not simply reduced opportunity for sex among those currently attending school (the so-called "incarceration effect" (Black, Devereux, and Salvanes 2008)). For men, education increased the likelihood of having more than one partner by 11.1 percentage points (p = 0.028), consistent with Kohler & Thornton (Kohler and Thornton 2011) but also increased literacy by 8.4 percentage points (p = 0.001), and discussion about HIV with others by 8.7 percentage points (p = 0.056). In interpreting these mediators, I caution that while sexual debut likely occurred prior to HIV infection, contemporary knowledge, attitudes, and behaviors were observed after all of the HIV-infected survey respondents had seroconverted, implying that the coefficients for secondary endpoints may capture behavioral responses downstream from HIV infection. However, for most secondary outcomes, these biases would be expected to run counter to my results: cohorts with higher HIV prevalence would be more likely to have ever tested for HIV, more likely to use condoms and to believe that it is acceptable for women to carry condoms, less likely to bear children, and more likely to talk about HIV. In fact, my results are in the opposite direction of these biases. Finally, although these outcomes are not monetary, it is possible that higher wage returns may be on the causal pathway, e.g. to economic independence of women. Prior studies suggest high wage returns to secondary schooling in Botswana (Siphambe 2000). I confirmed this relationship using the Botswana Labor Force Surveys, assessing the covariate-adjusted association between years of schooling and log hourly wage. One additional year of schooling was associated with a 20% (se=1.3) increase in wage in women; 12% in men (se=0.6) (Table 1.9, column 1). Using a Heckit model to adjust for truncation at zero, one additional year of schooling was associated with a 25% (se=1.3) increase in wage in women; 12% in men (se=0.6) (Table 1.9, column 2) (Wooldridge 2013). Future work will use the natural experiments described above to identify wage effects.

Table 1.6: *Empirical Strategy 1 (BAIS 2004 and 2008): Sexual Intercourse, Contraceptive Use, HIV Testing Behavior and Knowledge*

Coefficient on Years of Schooling	**(1)** **Female**	**(2)** **Male**
Dependent Variable		
Ever Had Intercourse (1=yes, 0=no)	-0.007	0.056*
	(0.027)	(0.030)
Observations	3,965	3,050
F-Statistic	8.6	9.6
Proportion Ever Had Intercourse, Pre-Reform	0.962	0.932
Age at First Intercourse	0.761***	0.065
	(0.261)	(0.209)
Observations	3,507	2,360
F-Statistic	11.9	10.4
Mean Age At First Intercourse, Pre-Reform	18.6	18.6
First sex ever: Did you use anything to protect yourself (eg, condom)? (1=yes, 0=no)	0.127***	0.055*
	(0.047)	(0.028)
Observations	3,582	2,458
F-Statistic	9.9	11.0
Proportion Protected, Pre-Reform	0.828	0.863
Ever Given Birth (1=yes, 0=no)	-0.158***	
	(0.057)	
Observations	3,644	
F-Statistic	10.0	
Proportion Ever Given Birth, Pre-Reform	0.728	
Have you ever been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	0.110*	0.120**
	(0.061)	(0.052)
Observations	3,793	2,922
F-Statistic	7.7	7.9
Proportion Ever Tested, Pre-Reform	0.720	0.573
Indicator for knowledge of at least one HIV prevention strategy (1=yes, 0=no)	-0.021	0.007
	(0.020)	(0.020)
Observations	3,791	2,919
F-Statistic	7.9	7.9
Proportion Yes on Indicator, Pre-Reform	0.949	0.952

Notes: All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. The instrument was defined as = 1 if YOB > 1980. The indicator for knowledge of HIV prevention strategies was defined as 1 if respondent could name at least one out of the six following HIV prevention strategies: condoms, fewer partners, mutually faithful relationship, abstinence, avoid injections with contaminated needles, and avoid blood transfusions. Those responding "don't know" to an HIV knowledge question were accounted for as incorrect. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table 1.7: *Empirical Strategy 1 (BAIS 2004 and 2008): HIV Misperceptions, Number of Partners, HIV Discussion and Attitudes, Literacy, and Labor Force Participation*

Coefficient on Years of Schooling	**(1)** **Female**	**(2)** **Male**
Dependent Variable		
Indicator for any misperceptions about HIV (1=yes, 0=no)	-0.062	-0.038
	(0.059)	(0.043)
Observations	3,782	2,915
F-Statistic	7.2	7.6
Proportion Yes on Indicator, Pre-Reform	0.564	0.603
Indicator for 2 or more sexual partners in the last 12 months (1=two or more, 0=one or zero)	0.044	0.111**
	(0.045)	(0.051)
Observations	3,658	2,877
F-Statistic	6.3	8.5
Proportion Two or More on Indicator, Pre-Reform	0.115	0.232
During the past 4 weeks, have you discussed HIV/AIDS with anyone? (1=yes, 0=no or not sure)	0.003	0.087*
	(0.060)	(0.046)
Observations	3,791	2,918
F-Statistic	7.7	8.2
Proportion Which Discussed HIV/AIDS, Pre-Reform	0.484	0.471
Do you think it should be acceptable for a woman to obtain male condoms? (1=yes, 0=no or not sure)	0.080**	0.096**
	(0.038)	(0.039)
Observations	3,832	2,957
F-Statistic	9.6	9.0
Proportion Yes on Outcome, Pre-Reform	0.933	0.875
Can you read and understand a letter / newspaper / bible? (1=easily, 0=no or with difficulty)	0.003	0.084***
	(0.036)	(0.023)
Observations	3,962	3,051
F-Statistic	8.7	9.7
Proportion Easily on Outcome, Pre-Reform	0.866	0.831
Labor Force Participation (1=yes, 0=no)	0.172**	0.048
	(0.076)	(0.039)
Observations	3,942	3,037
F-Statistic	8.4	9.3
Proportion Labor Force Participation, Pre-Reform	0.706	0.846

Notes: All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. The instrument was defined as = 1 if YOB > 1980. The indicator for any misperceptions was defined as 1 if respondent incorrectly answered any of the following four questions: whether HIV spreads via mosquitos, sharing a meal with an HIV+ person, due to witchcraft, and whether a healthy looking person can be HIV+. Those responding "don't know" to an HIV misconception question were accounted for as incorrect. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table 1.8: *Empirical Strategy 1 (BAIS 2004 and 2008): Age at First Intercourse*

	(1)	(2)	(3)	(4)
Dependent Variable	**Age at First Intercourse ≤ 16 Indicator**	**Age at First Intercourse ≤ 18 Indicator**	**Age at First Intercourse ≤ 20 Indicator**	**Age at First Intercourse ≤ 22 Indicator**
Coefficient on Reform Indicator		*Female Subsample*		
Female	-0.091***	-0.116***	-0.083***	-0.018
	(0.029)	(0.036)	(0.025)	(0.012)
Observations	3,507	3,507	3,507	3,507
R-squared	0.041	0.104	0.073	0.053
Mean Dependent Variable, Pre-Reform	0.187	0.544	0.850	0.963
		Male Subsample		
	0.027	-0.039	-0.046	0.001
	(0.041)	(0.044)	(0.032)	(0.020)
Observations	2,360	2,360	2,360	2,360
R-squared	0.046	0.105	0.108	0.080
Mean Dependent Variable, Pre-Reform	0.228	0.491	0.789	0.931
		Both Sexes		
	-0.044*	-0.085***	-0.068***	-0.011
	(0.024)	(0.028)	(0.020)	(0.011)
Observations	5,867	5,867	5,867	5,867
R-squared	0.047	0.105	0.092	0.071
Mean Dependent Variable, Pre-Reform	0.205	0.522	0.824	0.949

Notes: All regressions are OLS models, which include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0·01, ** p<0·05, * p<0·1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table 1.9: *Earnings Regressions (Labor Force Surveys 1995/96 and 2005/06)*

Dependent Variable: Log(Hourly Wage)	**(1)** **OLS**	**(2)** **Heckit**
Coefficient on Years of Schooling		
Female Wage Earners	0.198***	
	(0.013)	
Female Whole Sample		0.250***
		(0.013)
Observations	1,139	5,363
R-squared	0.496	
Coefficient on Years of Schooling		
Male Wage Earners	0.123***	
	(0.006)	
Male Whole Sample		0.122***
		(0.006)
Observations	1,272	4,711
R-squared	0.584	

Notes: Regression 1 is an OLS model. Regression 2 is a Heckit model using Stata's heckman command, where the likelihood of a wage being observed is a function of marital status, age and education. Both models control for a continuous quadratic term in age, an indicator for birthplace, and an indicator for survey wave. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and born in or after 1975. Observations with wages above the 95th percentile or below the 5th percentile were dropped as outliers. No weights were used. Source: Botswana Labor Force Survey 1995/96 and 2005/06.

1.8 Discussion

Using an education policy reform as a natural experiment, I find that secondary schooling has a large protective effect against risk of HIV infection in Botswana. Effects were particularly large for women and were consistent through a wide array of robustness checks. The IV estimates for HIV are somewhat larger, but generally consistent with the strong negative associations between secondary schooling and HIV infection risk in OLS regressions. One explanation for why the 2SLS results are larger than the OLS results could be that unobserved factors, such as personal charisma or ability, may be positively associated with both educational attainment and HIV risk, thereby reducing the magnitude of the OLS coefficient. Another explanation is measurement error, although it is unlikely that noise in reported years of schooling is large enough to account for the difference between OLS and IV coefficients. A third explanation is that – as with all IV estimates – the causal effects are "local" to the subpopulation of compliers, e.g., those induced to increase schooling because of the reform. This subpopulation consists of persons who, in the absence of the reform, would have dropped out after ninth grade – a group likely to be at high risk for HIV.

My effect estimates for a single year of schooling are large; however, there is reason to believe that the later years of secondary school are particularly protective against HIV infection risk. The OLS regressions provided suggestive evidence that the effect of schooling on HIV risk may be non-monotonic in Botswana: an additional year of schooling was associated with slightly *increased* HIV risk from years of schooling 0 to 9; but with large reductions in HIV risk in years 10-13+. Indeed, the education gradient in HIV infection is highly non-linear in five other sub-Saharan African countries with sharp declines in HIV infection risk above 9 or 10 years of schooling (de Walque 2009, Fortson 2008). These results are also consistent with findings from HIV prevention interventions targeting adolescents, including in school settings (Coyle et al. 2004, Coyle et al. 2006, Tortolero et al. 2010), that have been particularly effective at reducing HIV infection risk (DiClemente et al. 2009, Jemmott et al. 2005, Jemmott, Jemmott, and Fong 2010, Kirby 2008, Koniak-Griffin et al. 2003, Shrier et al. 2001, Sikkema et al. 2005, Tortolero et al. 2010, Villarruel, Jemmott, and Jemmott 2006). Multiple countervailing pathways from education to HIV risk may be at work. Participation in late primary and early secondary school may increase social (and sexual) network size but confer little in the way of economic opportunity or cognitive skills to navigate a complex risk environment. Although I do not have a natural experiment for primary schooling, it is quite likely that results would differ substantially. The later grades of secondary school may be a critical exposure period in the determination of HIV risk (Kuh 2003).

Additional years of schooling had a causal effect on some proximate risk factors for HIV, but not others, providing insights on potential mechanisms. Education increased condom use and improved norms regarding women carrying condoms, suggesting more widespread adoption of this particular prevention technology among the better educated, a phenomenon suggested in other studies (Cutler and Lleras-Muney 2010, de Walque 2009, Hargreaves et al. 2008). Education increased HIV testing for both men and women and led to increases in the proportion of men who reported talking with others about HIV, suggesting increased openness about HIV and demand for knowledge about one's own HIV status. Additional schooling led women (but not men) to delay sexual debut and to delay childbearing. The reduction in childbearing is generally consistent with reduced unprotected sexual activity and lower risk for HIV, and is also of interest in its own right, particularly when paired with the large increase in labor force participation. In contrast to work by Case and Paxson (Case and Paxson 2013), I find the exogenous changes in education led to shorter (not longer) durations of pre-marital sexual activity, which may have reduced exposure to HIV.

Interestingly, education had no effect on abstinence (measured at the time of the surveys) or number of partners for women, and increased numbers of partners for men

– a finding similar to other literature (Fortson 2008, Kohler and Thornton 2011). Although I have limited data on partner characteristics, I find no evidence that education reduced participation in age-disparate relationships – "sugar daddies/sugar mamas", thought to be a driver of HIV risk (Harling et al. 2014, Leclerc-Madlala 2008, Longfield et al. 2004, Luke 2005). Whereas education neither reduced sexual activity, nor reduced numbers of sex partners, nor lead people to select younger partners – it did appear to delay sex for women and to reduce HIV risk *within* relationships through increased awareness of HIV status, communication about HIV, and normalization of condom use. These findings contribute to our understanding of the margins on which Batswana with additional schooling successfully adapted to reduce exposure to HIV in a highly endemic setting.

What was it about education that led to these changes in behavior and reductions in HIV infection risk? Perhaps the most obvious hypothesis is that schooling might provide information about HIV that enables people to make safer decisions. In fact, gains in education induced by the reform had zero causal effect on HIV knowledge and misconceptions. Although counterintuitive, this result is not surprising given that the reform cohorts in my study completed secondary school in the 1990s and early 2000s, before Botswana launched a formal HIV curriculum in schools. Many resources have been devoted to HIV-specific education programs, and indeed HIV knowledge was observed to be high in Botswana. However, my results suggest that the effect of secondary schooling on HIV risk is not attributable to knowledge acquisition, but rather to other factors. Scholars have long argued that "knowledge is not enough" to prevent HIV infection (Levinsohn, Dinkelman, and Majelantle 2006). My results indicate that secondary education provides critical enabling factors that allow knowledge to be utilized and enable people to avoid HIV infection.

Although education had no effect on HIV-specific knowledge, secondary schooling did have large causal effects on other factors that may have mediated the effect of schooling on HIV risk. These pathways differed for men and women. For women, the additional schooling induced by the reform had very large effects on labor market participation, a consequence of the reform that has also been reported elsewhere (Borkum 2009). In the BAIS data, the reform caused over half of those women who would have otherwise been out of the labor force to seek employment and increased wages by nearly a quarter (per year of schooling induced by the reform). These changes in economic opportunity may have enabled women to make explicit choices to reduce HIV risk, e.g. by increasing bargaining power and financial independence for women within relationships (Baird et al. 2012, Pollak 2005). Indeed, I find evidence that education changed norms for both women and men about the acceptability of women carrying condoms, and led to increases in condom use and HIV testing. In addition to economic empowerment, schooling may have also led women to think differently about their future, changing expectations about whether they would have their own career or be a homemaker, and changing childbearing preferences. Indeed, I observe a large delay in sexual debut and a marked reduction in childbearing for women affected by the reform. Interestingly, I observe a rightward shift in the distribution of age at sexual debut not just from 16 to 17, the age when most girls would have been in grade ten, but through age 20, suggesting that education did not just limit opportunities for sex among school-going youth, but actually changed preferences towards later sexual debut.

The effects of schooling on women's labor market participation and childbearing have been observed in other settings without hyper-endemic HIV (Ainsworth, Beegle, and Nyamete 1996, Martin 1995, McCrary and Royer 2011, Monstad, Propper, and Salvanes 2008, Osili and Long 2008, Sackey 2005), and schooling has been cited as a key policy lever in spurring the demographic transition from high to low childbearing (Becker, Cinnirella, and Woessmann 2010, Bloom and Canning 2004, Bloom et al. 2009, Caldwell 1980). In Botswana, investments in education appear to have had the added benefit of reducing HIV infection. Whether the observed reductions in HIV risk are unintentional consequences of increased utilization of condoms as birth control, or a result of explicit decisions by economically empowered and/or more forward-looking women to reduce their exposure to HIV, cannot be ascertained. However, the benefits of secondary school in reducing HIV risk appear to be closely related to effects of education on the economic and reproductive empowerment of women in Botswana.

Although the additional education induced by the reform had relatively little economic impact for men (in terms of labor force participation), I find evidence that extra years of schooling may have had impacts on cognitive skills, as suggested by very large increases in literacy. The proportion of men reporting that they could not read and understand a letter or newspaper with ease dropped by nearly half per year of schooling induced by the reform. These skills may have improved men's ability to use information about HIV in making complex decisions in their lives (Cutler and Lleras-Muney 2010). Indeed I find evidence that men are more likely to discuss HIV with others and, similarly to women, to test for HIV, use condoms, and report that it is acceptable for women to carry condoms. Although knowledge about HIV prevention methods did not change, men may have acquired additional skills to utilize that knowledge in their lives.

In summary, I draw three conclusions from this discussion of mechanisms. First, education had no effect on HIV knowledge, but did affect norms and behaviors. Second, observed changes in behavior occurred on the margins of risk reduction *within* relationships and delayed sexual debut for women. I found no evidence of partner reduction, abstinence after debut, or selection of less risky partners as pathways. Third, education led to fundamental changes in market labor supply for women, which may have empowered women to reduce exposure to HIV due to increased bargaining power, financial independence, future-orientation, or changed childbearing preferences. The ability to chart out causal pathways in this degree of detail is rare and a testament to the data collected in the BAIS surveys. Finally, the 1996 reform affected a number of proximate HIV risk factors (e.g., condom use, sexual debut) and outcomes correlated with unprotected sex (e.g., childbearing) generates added confidence in my results.

This study has some limitations. First, consent rates were imperfect, and migration or mortality could have influenced the composition of the study sample. However, neither consent rates nor birth cohort sizes varied systematically with exposure to the reform, suggesting that any bias from these sources is minimal. I also imputed HIV status for people without valid HIV test result and find similar results using the full sample. Second, my analysis relies on the assumption that HIV infection does not cause people to stay in secondary school or drop out. Infection rates are likely very low prior to grade ten, with only 10.1% of women and 14.9% of men having their sexual debut before age 16 (BAIS 2008). The vast majority of people infected with HIV would be asymptomatic during the period of their schooling making reverse causality unlikely. In addition, I focus on birth cohorts whose childhood occurred prior to the advent of pediatric antiretroviral therapy (ART). Any children infected during birth or breastfeeding would almost certainly have died prior to secondary school and would not be in the sample. Third, I only observe people through age 32 years in the BAIS datasets. I cannot know whether I are measuring HIV infections truly averted or delayed. However, this is a common limitation of prevention studies. In spite of this limitation, my analysis of cumulative incidence captures much longer follow-up than most RCTs, which observe incidence over a shorter, e.g., 3yr (Pettifor et al. 2015), horizon.

Fourth, by exploiting exogenous variation in schooling, I avoided issues of self-selection of high-risk individuals into more (or less) schooling, and thereby control for such unobserved confounders as: socioeconomic status, risk aversion, future-orientation, self-confidence, etc. However, my analysis nevertheless relies on the assumption that conditional on covariates – there were no other cohort-specific effects that would have affected HIV risk aside from exposure to the reform. There are many reasons why HIV risk might change across birth cohorts but the likely candidates – infection rates among sexual partners, access to HIV treatment, changes in prevention programming – are phenomena that are either gradual over time (changes in the epidemic context) or, if they are sudden, affect people of many different ages (e.g., a national prevention campaign): in both cases, these phenomena would result in gradual changes in HIV infection across birth cohorts. I control for such changes in risk across birth cohorts using a linear (or quadratic) trend, which picks up all observed and unobserved factors that change smoothly (linearly, quadratically) across birth cohorts. The validity of my natural experiment would be jeopardized only if there were unobserved factors that led to a discontinuous change in HIV risk for the specific birth cohorts affected by the policy reform. One example would be an HIV prevention program targeted to a specific grade in school. However, Botswana's HIV education programming was not in place in 1996 (Government of Botswana 1995, Tsheko 2012). In robustness checks, an alternative difference-in-differences identification strategy yielded similar results, lending support for this assumption. Finally, as with all infectious diseases, there could be spillover effects beyond the individuals directly affected by the reform, which would change the interpretation of my estimates; however, given that people have sexual relationships across birth cohorts, these spillovers would be expected to have a smooth impact on HIV infection across birth cohorts and would not bias my estimates, which are based on a discontinuous change in exposure to the reform.

Many studies have reported correlations between schooling and HIV infection risk (Barnighausen et al. 2007, Dallabetta et al. 1993, Fylkesnes et al. 2001, Hargreaves et al. 2008, Kirunga and Ntozi 1997). This study is among the first to use a natural experiment to assess the causal relationship between schooling and HIV infection. Expanding access to secondary school had a large protective effect against HIV risk in

Botswana. In settings with large, generalized HIV epidemics, estimates of the returns to secondary schooling may be underestimated due to the exclusion of health benefits.

Chapter 2

Children's Education and Parental Old Age Survival – Quasi-experimental Evidence on the Intergenerational Effects of Human Capital Investment[1]

2.1 Introduction

Intergenerational wealth transfers have increasingly been recognized as a critical mechanism to understand well-being and wealth accumulation within families and societies (Lee and Mason 2011). While a large literature has documented the large resource flows from parents to children (Folbre 1994), a growing literature from Asia (Cai, Giles, and Meng 2006, Frankenberg, Lillard, and Willis 2002, Giles, Wang, and Zhao 2011, Knodel et al. 2000), Latin America (Bravo 2006, Saad 2005), and sub-Saharan Africa (Adamchak et al. 1991, Kohler et al. 2012) suggests that intergenerational transfers have started to reverse in aging low-income societies where governmental old age support programs are generally limited (Bloom, Canning, and Fink 2011).

Anecdotal evidence from developing countries suggests that altruistically motivated children are generally expected to share their resources with parents (Chou 2008, Lillard and Willis 1997, Robert and Stark 1985). The past decades have brought a massive increase in educational opportunities for younger cohorts. The extent to which this younger generation shares their increased knowledge (Kuziemko 2014), skills (Evans and Becker 2009), and available resources (Kuziemko and Ferrie 2014) with their parents remains largely unknown (Cameron and Cobb-Clark 2006). Previous studies on the role of children's human capital in parental old age survival have generally been

[1] Co-authored with Günther Fink. This essay has been submitted in part for publication.

descriptive in nature, using either cross-sectional (Zimmer, Hermalin, and Lin 2002, Yahirun, Sheehan, and Hayward 2016, Sabater and Graham 2016) or longitudinal data (De Neve and Harling 2017, Friedman and Mare 2014, Torssander 2013). While these studies generally find positive correlations between children's human capital and parental survival (De Neve and Kawachi 2017), the extent to which these results can be given causal interpretation seems highly limited. Given the large number of potential factors (such as genetic traits) determining both (Krapohl et al. 2014), these correlations are hard to interpret, and likely do not identify causal effects of educational opportunities provided to children. A recent paper exploring increases in education as a proxy for life-time earnings finds no evidence of a causal effect of children's resources on parental wellbeing (Lundborg and Majlesi 2015).

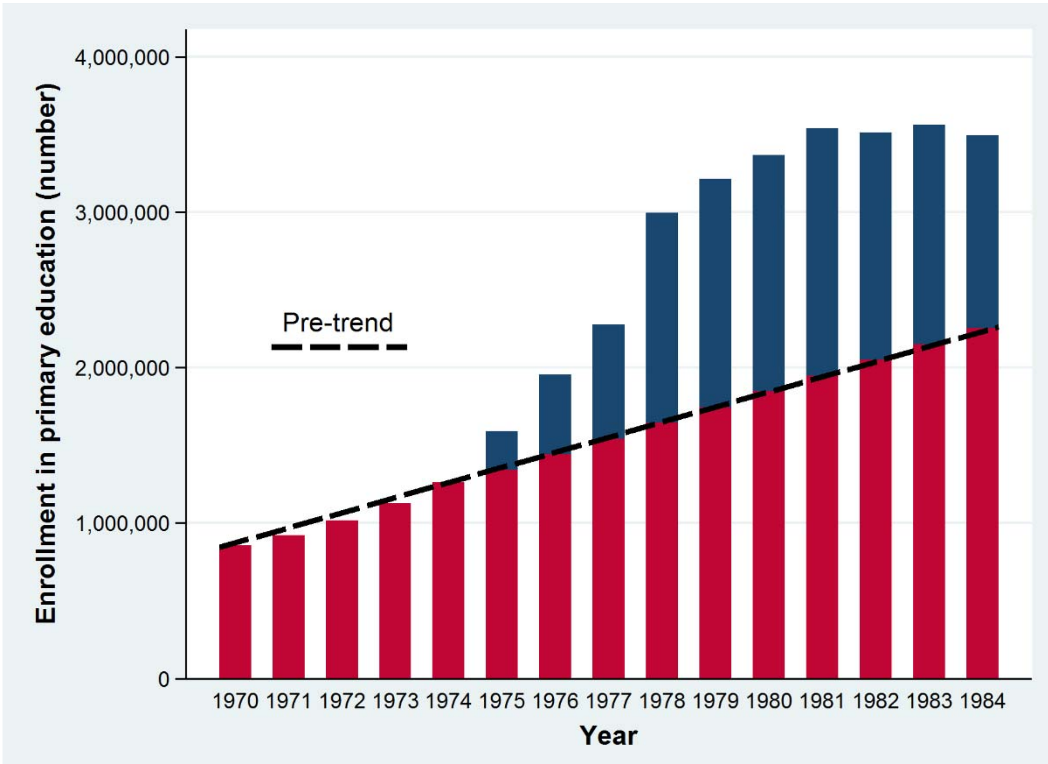
In this study, I explore one of the largest schooling reforms in low- and middle-income countries to assess the impact of increased educational investment in younger generations on parental survival. In 1974, Tanzania launched a major Universal Primary Education initiative (Oketch and Rolleston 2007). Between 1974 and 1979, total primary school enrollment in the country increased from 1,260,103 to 3,211,568 (UNESCO Institute for Statistics 2015) (Figure 2.1). At the cohort level, the reform induced 1.9 additional years of educational attainment relative to the attainment the affected cohorts would have achieved if previous trends had continued (Figure B1 in the supplementary materials; see Table 2.2 for regression results). I use this discontinuity in educational attainment to identify the causal effect of children's education on parental survival.

2.2 Policy Reform and Study Context

Tanzania is located in East-Africa and one of the poorest countries in the world. Tanzania's GNI per capita was 920 US$ in 2014 with about half of the population living on less than $1.90 a day (2011 PPP). Life expectancy at birth grew from 43.7 in 1960 to 64.3 years in 2013 (The World Bank 2015).

Tanzania's public education system is based on the "7-4-2-3" system, including seven years of primary school, four years of ordinary level secondary school, two years of advanced level secondary school, and three years of university education. In the 1970s, the government of Tanzania strived to achieve Universal Primary Education (UPE) and embarked on a large-scale national campaign for universal access to primary education, of all children of school going age (Nyerere 1967). This effort culminated in 1974, when the government of Tanzania announced the Musoma Resolution on UPE (Agwanda and Amani 2014). As a result, total primary school enrollment in the country increased from 1,260,103 to 3,211,568 between 1974 and 1979 (UNESCO Institute for Statistics 2015) (Figure 2.1). Primary school completion was considered the final stage of education (Sabates, Westbrook, and Hernandez-Fernandez 2011). Previous studies have found that the policy reform led to a large increase in primary schooling attainment (Omari 1983, Sabates, Westbrook, and Hernandez-Fernandez 2012).

Fig. 2.1: *Enrollment in Primary Education in Tanzania, by Year*



Notes: Figure shows enrollment numbers in primary education in Tanzania (both sexes) by year. In 1974, Tanzania launched a major initiative to achieve Universal Primary Education within five years. Between 1974 and 1979, total primary school enrollment in the country increased from 1,260,103 to 3,211,568. The sample includes individuals officially registered in a primary school, regardless of age. Source: administrative data from UNESCO.

2.3 Data Description

2.3.1 Study Population

The study population included all citizens of Tanzania, born in Tanzania, and residing in the country in 1988 and 2002. Respondents younger than 15 years old were excluded because they would not have had the opportunity to complete their primary education (Parliament of Tanzania 1977). Respondents born prior to 1945 were excluded because these cohorts were mostly trained in the pre-independence era. Those not born in Tanzania were excluded because they would not have been exposed to the schooling intervention if they migrated in adulthood.

2.3.2 Sampling Strategy

Data for this study were obtained from the United Republic of Tanzania Population Census 1988 and Population and Housing Census 2002 through the Integrated Public Use Microdata Series (IPUMS 2014). Different from standard Censuses, Tanzania did not enumerate all households, but rather selected a random sample of about 20% (census 1988) and 15% (Census 2002) of Census Enumeration Areas. The Integrated Public Use Microdata Series (IPUMS) drew a systematic 1-in-2 sample from the original 20% sample (Census 1988) and a two-thirds subsample from the original 15% sample (census 2002) to provide researchers with a 10% "real" sample. All individuals in the selected areas were interviewed with the exception of diplomats and their families. Parental survival was included in the survey for all respondents while educational attainment was included for respondents five years old and over. Data on demographics, province of birth, years of schooling, and *maternal* survival was available for over 99.99% of eligible respondents, yielding a total of 4,977,711 individuals in the

1988 and 2002 Census. Data on demographics, province of birth, years of schooling, and *paternal* survival was available for over 99.99% of eligible respondents in the 1988 Census' universe, yielding 1,896,911 individuals in the 1988 Census.

2.3.3 Measurement of Exposures and Endpoints

I focus on the child as unit of analysis with his/her schooling as exposure, regardless of a child's age. Data on educational attainment, age, gender, citizenship, province of birth, marital status, area of residence, labor force participation, literacy, basic household utilities, parental mortality status and co-residence were extracted from the Census datasets. Year of birth was calculated based on the age collected in the Census. The key independent variable of interest was the child's total years of schooling completed. My main outcome of interest was "mortality status of mother" and "mortality status of father" indicating whether a respondent's biological mother and father were still living at the time of the Census. As an additional outcome of interest, I created a variable for "mortality status of spouse" based on a respondent's marital status (i.e., widowed vs. married, separated, or divorced). Pointer variables provided by the IPUMS system also make it possible to construct individual-level variables representing the characteristics of co-resident family members (e.g., parental age and years of schooling). Data on date of death, cause of death, and education of those who had died were unavailable.

2.4 Empirical Approach

2.4.1 Causal Effect of Children's Education on Parental Survival

I explore the quasi-experimental variation generated by the 1974 education reform. In Tanzania, children are expected to start primary school at age seven years. Therefore, individuals born in 1967 or later would have entered primary school in 1975 and would have been fully exposed to the reform. Individuals born between 1960 and 1966 were between 8 and 15 years old when the reform started and thus should have been at least partially affected by the reform – these cohorts were treated as "partially exposed" in my analysis.

In a first step, I estimate the causal impact of the reform on educational attainment using multivariable interrupted time series modeling. I also show the additional educational attainment attributable to the reform (deviations from pre-existing time trends) graphically. Second, I assessed the intention-to-treat (ITT) effect of being in a reform cohort on parental survival. Third, I estimated 2SLS (instrumental variables) regression models, using exposure to the reform as an instrument for total years of schooling. In all models, I controlled flexibly for an indicator for five-year age groups to account for the non-linear pattern in mortality across ages as well as age heaping in reported ages[2]. In regressions for maternal survival, I additionally controlled for an indicator for Census year since maternal survival status was available for two Census years. Exposure to the reform was modeled as an intercept shift for cohorts affected by the reform. I interpret my instrumental variable estimates as specific to the subpopulation who complied with their treatment assignment – i.e., children who increased their schooling because of the reform (Angrist, Imbens, and Rubin 1996).

[2] I also included an indicator for heap year and included indicators for province of birth.

As a benchmark for my causal analysis, I also assessed the naïve association between years of schooling and parental mortality status. I assessed the crude relationship graphically and then adjusted for time trends and other covariates available in the Census data. I first estimated models for women and men separately, and then on the pooled sample. When pooling sexes, I included indicators for sex and the interactions of sex with all other covariates.

2.4.2 Robustness Checks

I conducted a wide range of sensitivity analyses. The assumption that underlying trends are approximately linear is more plausible when few cohorts are included. In my first robustness check, I restricted the sample to households just around the cutoff (+/- 5 years). I also included alternative specifications of age, including a linear, quadratic, and quartic term in age to allow for non-linearities in underlying cohort trends. To assess whether intergenerational correlation in longevity and educational attainment could explain my main results (i.e., due to changing trends in parental age and education), I graphically assessed the distribution of age and educational attainment among surviving parents by birth cohort. I assessed the relationship between educational attainment and parental mortality in 2002 in multivariable logistic regression models, controlling for predicted maternal age and education from the Tanzania Census of 1988. In analytical inference, the use of sample weights is subject to controversy. I added sample weights to my main model as an additional robustness check. I also used alternative definitions for exposure to the reform and analytical sample, excluded birth cohorts partially exposed to the reform, and included the full sample of children in my analysis.

The availability of two Census years with data on maternal mortality status for Tanzania allowed me to generate variation in the instrument within a single age group and control simultaneously for age and period effects in my regressions for maternal survival. It also allowed me to, implicitly, control for birth cohort effects. To generate added confidence that my results are not confounded by other policy changes, I estimated 2SLS difference-in-differences models exploiting the fact that the policy reform had a larger impact in some provinces of birth than others, based on the share of respondents with no schooling in the pre-reform period (Figure B2 in the supplementary materials displays school enrollment by province for each birth cohort).

Differential migration or mortality of children could have influenced the composition of the study sample. However, children's birth cohort sizes did not vary systematically with exposure to the reform. It is also possible that the parents of some respondents had deceased prior to the age when a respondent would have entered primary school. Parental mortality rates, however, are low prior to grade one. More importantly, since I are assessing the effect of education on parental survival on a risk difference scale, my approach is robust to the existence of prevalent parental mortality by grade one, so long as prevalence was smooth across birth cohorts.

2.4.3 Age-specific Survival

One of the limitations of the Census data is that it does not allow me to directly control for the age of deceased parents, since time of death of parents was not available in the Census data. To overcome this limitation as well as to quantify the protective impact of child education across parental age groups, I analyze spousal survival rates in households where at least one parent is alive and living in the same household as their child (33% in the 1988 Census and 33% in the 2002 Census). At the population level, average spousal ages are relatively closely aligned (40% of cohorts ages 30 – 40 years old were within +/- 5 years of their spouse), so that I can analyze the probability of survival by approximate age group. I estimated 2SLS models assessing the effect of additional children's schooling on spousal survival among mothers co-residing with their children. I estimated separate 2SLS models for five-year maternal age categories, controlling flexibly for time trends in addition to other covariates available in the Census data.

2.4.4 Survival Rates

I calculated cohort-specific survival rates between the 1980s and 2000s under the assumption that net migration is zero (migration is limited in Tanzania at older ages (Agwanda and Amani 2014). Although I do not observe the same individuals over time, I calculated a 14-year survival rate using the 1988 and 2002 Census, as the number of people alive in 2002 (N_{02}) divided by the number of people alive in 1988 (N_{88}) for each birth cohort. I assessed the relationship between survival rates and parental exposure to the policy reform in two ways. First, I ignored the exact relationship to the parental generation, and analyzed whether or not the household (in some way) benefited from the reform in the 1988 and 2002 Census. To do so, I used descriptive multivariable OLS linear probability models and regressed cohort survival rates on a continuous variable for the likelihood of respondents with children benefitting from the reform based on the year of birth of their children. Second, since the child of a respondent is likely to be a grandchild of my cohort of interest in previous analyses (i.e., a respondent's parent), I also used the 1988 data to compute the percentage of respondents having a reform-benefitting child for each cohort and regressed survival rates in the 2002 Census on this variable. I controlled for time trends and cohort educational attainment in addition to other covariates.

2.4.5 Causal Pathways from Children's Education to Parental Survival

To investigate the causal pathways from education to parental survival, I assessed the effect of the reform on potential mediators. I ran my main 2SLS model for secondary outcomes, including labor force participation, literacy, marital status, spousal schooling, and access to basic household utilities. I ran my main OLS model assessing the relationship between a child's schooling and parental survival, but iteratively adding potential mediators as controls to see how much of the education effect on parental survival remains after shutting down these pathways. To further assess the likelihood of non-income pathways, I also ran my main 2SLS model for the main outcome, parental survival, but adding labor force participation as an additional control.

This study was considered exempt from full review by the Harvard T.H. Chan School of Public Health Institutional Review Board since it is based on an anonymous public use data set with no identifiable information on survey participants. Stata (version 14.0, Statacorp, College Station, Texas) was used for all statistical analyses.

2.5 Results

In my analytical sample, the average child with data on parental mortality status was 28 years old (Table 2.1); the average age of surviving parents co-residing with their children was 47 (mothers) and 55 (fathers) years.

Table 2.1: *Summary Statistics*

Census Year	1988 (N=902,568)	2002 (N=1,804,198)	Both Years (N=2,706,766)
Individual			
Mother Alive, %	83.8	77.9	80.0
Father Alive, %	n/a	62.5	n/a
Age, Mean (SD)	25.8 (7.9)	29.8 (11.1)	28.4 (10.1)
Years of Schooling, Mean (SD)	4.7 (3.3)	5.2 (3.4)	5.0 (3.4)
Has At Least One Year of Schooling, %	72.1	75.9	74.5
Has At Least Seven Years of Schooling, %	52.7	60.1	57.4
Labor Force Participation, %	82.9	78.0	79.7
Literacy, %	72.4	74.3	73.6
Household			
Number of Own Family Members in Houshold, Mean (SD)	6.2 (4.1)	5.3 (3.4)	5.6 (3.7)
Number of Mothers in Household, Mean (SD)	1.1 (0.8)	1.0 (0.7)	1.0 (0.7)
Number of Fathers in Household, Mean (SD)	0.7 (0.6)	0.7 (0.6)	0.7 (0.6)
Any Death(s) in Household Last Year, %	7.7	6.9	7.1
Urban, %	40.2	27.4	31.7
Access to Toilet or Latrine, %	89.4	91.3	90.6
Access to Piped (Running) Water, %	35.1	34.4	34.6
Access to Electricity, %	8.1	11.3	10.1

Notes: Sample includes Tanzanian citizens, born in Tanzania after 1945, and ages at least 15 years old. In the 1988 census, paternal mortality status is not available; and regional and district headquarters are considered urban, whereas villages are considered rural. In the 2002 census, the urban-rural status is stated explicitly. Sample weights used as provided. Source: Tanzania Census 1988 and 2002.

2.5.1 Effect of the 1974 Education Reform on Years of Schooling Completed

The reform increased average years of schooling by 1.9 years ($p < 0.001$) for cohorts fully affected by the reform (Table 2.2). Figure B1 shows completed years of schooling by year of birth. Years of schooling increased dramatically for cohorts born in 1967 or later. Children born between 1960 and 1966 were partially exposed to the reform. Due to grade repetition and/or late entry into school, some respondents born before 1960 likely were also affected by the reform.

Table 2.2: *First Stage Results: Effect of the Policy Reform on Schooling*

Dependent Variable	**Schooling (Years)**			
Age Specification	**Age**	**Age, Age^2**	**Age, Age^2, Age^3, Age^4**	**Five-Year Age Group Indicator**
Predictor				
Coefficient on Partial Reform Exposure Indicator	1.620***	1.377***	1.383***	1.533***
	(0.007)	(0.008)	(0.008)	(0.008)
Coefficient on Full Reform Exposure Indicator	1.992***	1.784***	1.801***	1.865***
	(0.010)	(0.010)	(0.010)	(0.010)
Observations	2,706,766	2,706,766	2,706,766	2,706,766

Notes: OLS models. In addition to controlling for f(age), all regressions include an indicator for heap year, indicator for census year, indicators for province of birth, indicator for sex, and the interactions of each covariate with sex. Full exposure to the reform was defined as 1 if year of birth >= 1967, zero otherwise. Partial exposure to the reform was defined as 1 if year of birth 1960 - 1966, zero otherwise. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights were used. Sample includes Tanzanian citizens, born in Tanzania after 1945, and ages at least 15 years old. Source: Tanzania Census 1988 and 2002.

2.5.2 Descriptive Association between Children's Education and Parental Survival

The crude association between additional children's education and parental survival was non-linear. I find strongest protective effects for primary school, with protective effects levelling off after lower secondary school (Figure B3 in the supplementary materials). The association between children's education and parental survival persisted in multivariable OLS regression models (Model A, Table 2.3). Each additional year of schooling among children was associated with an increase of 1.0 percentage points ($p < 0.001$) in both maternal and paternal survival in the pooled sample, while women's schooling was associated with an increase of 1.0 percentage points ($p < 0.001$) and men's schooling with an increase of 0.8 percentage points ($p < 0.001$).

2.5.3 Causal Effect of Children's Education on Parental Survival

Table B3 in the supplementary materials presents ITT results, in which parental survival was regressed directly on the instrument and covariates. Mothers of children fully exposed to the reform were 4.7 percentage points ($p < 0.001$) more likely to be alive at the time of the Census; whereas fathers of children fully exposed to the reform were 2.9 percentage points ($p < 0.001$) more likely to be alive. Table 2.3 shows 2SLS results (Model B). Each additional year of schooling obtained by a daughter decreased the probability of maternal death by the time of the Census by 2.5 percentage points ($p < 0.001$), while each additional year of schooling attained by a son decreased the probability of maternal death by 3.9 percentage points ($p < 0.001$). For fathers, the probability of death decreased by 2.2 percentage points for female descendant education ($p = 0.054$), and by 10.8 percentage points for male descendant education ($p < 0.001$). These differences are sizeable relative to the absolute mortality risk: on average, only 26% and 45% of children born post-reform without education had lost their mothers and fathers, respectively, in 2002 (Table B4).

Table 2.3: *OLS and 2SLS Results: Children's Education and Parental Survival*

Dependent Variable	**Mother Alive (1=yes, 0=no)**			**Father Alive (1=yes, 0=no)**		
Subsample	**Female**	**Male**	**Both Sexes**	**Female**	**Male**	**Both Sexes**
A: OLS Model						
Schooling (Years)	0.010***	0.008***	0.010***	0.010***	0.008***	0.009***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Probability Dependent Variable, Pre-Reform	0.732	0.758	0.745	0.390	0.425	0.407
B: 2SLS Model						
Schooling (Years)	0.025***	0.039***	0.031***	0.022*	0.108***	0.047***
	(0.001)	(0.001)	(0.001)	(0.012)	(0.034)	(0.012)
Observations	1,442,732	1,264,034	2,706,766	958,122	846,076	1,804,198

Notes: Panel A shows regression results from an OLS model controlling for indicators for five-year age group, an indicator for census year (in the regressions for Mother Alive), indicator for heap year, and indicators for province of birth. Mother Alive captures maternal survival status at the time of the 1988 and 2002 census. Father Alive captures paternal survival status at the time of the 2002 census. Panel B shows regression results from a 2SLS model, in which exposure to the reform was used as an instrument for years of schooling. The instruments were an indicator for partial exposure to the reform (year of birth 1960 - 1966) and an indicator for full exposure to the reform (year of birth >= 1967). Regressions for the subsample with both sexes additionally control for an indicator for sex and interactions of each covariate with sex. *F*-statistics were 16,490 (both census years) and 81 (census 2002) for the female subsample; 6,005 (both census years) and 16 (census 2002) for the male subsample; and 20,342 (both census years) and 88 (census 2002) in the pooled sample. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Tanzania, born in Tanzania after 1945, and ages at least 15 years old. No weights were used. Source: Tanzania Census 1988 and 2002.

2.5.4 Sensitivity Analyses

In general, results were not sensitive to different modelling strategies for age including a linear, quadratic, or quartic term in age, the use of sampling weights, nor to alternative sample specifications including restricting my analytical sample to cohorts born +/- 5 years around the policy reform cut-off, the full sample of birth cohorts, and when excluding cohorts that were partially affected by the reform (Table 2.4 and Tables B1 – B2 in the supplementary materials). Parental age and education trends were smooth across children's birth cohorts (Figure 2.4 in the supplementary materials). In my difference-in-differences models, the reform had a larger effect for children in provinces where a higher share of the population had no schooling pre-reform.

Table 2.4: *Robustness Checks: Assessing the Robustness of the 2SLS Results to a Quadratic in Age, Sampling Weights, Alternative Sample Definitions, and Alternative Identification Strategy*

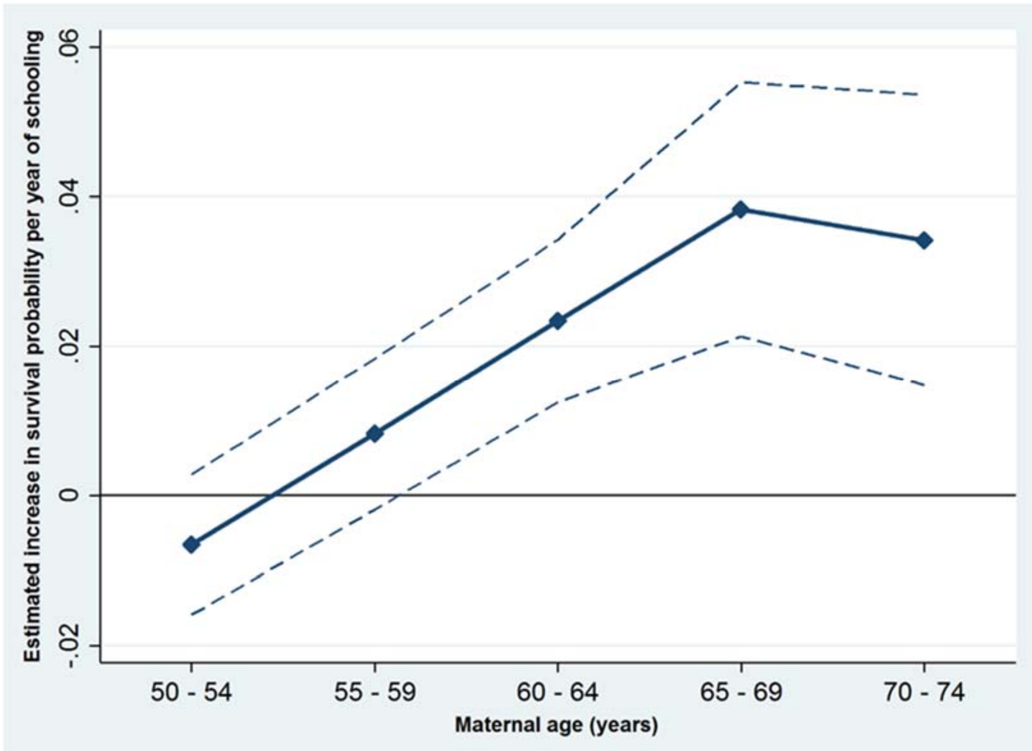
Specification	**2SLS, Quadratic in Age**	**2SLS, Sampling Weights**	**2SLS, Full Sample of Cohorts**	**2SLS, Cohorts Around Cut-off**	**2SLS, Excluding Partially Exposed**	**2SLS, Difference-in-Differences**
Exposure	**Schooling**	**Schooling**	**Schooling**	**Schooling**	**Schooling**	**Schooling**
Dependent Variable: Mother Alive	0.024***	0.028***	0.026***	0.114***	0.021***	0.042***
	(0.001)	(0.001)	(0.001)	(0.010)	(0.001)	(0.002)
Observations	2,706,766	2,706,766	3,251,376	1,138,634	2,221,198	2,706,766
Dependent Variable: Father Alive	0.067***	0.034**	0.060***	0.104***	n/a	0.031***
	(0.005)	(0.015)	(0.014)	(0.016)	n/a	(0.002)
Observations	1,804,198	1,804,198	2,035,469	535,329	n/a	1,804,198

Notes: Table presents robustness checks for the 3.1 and 4.7 percentage point reductions in maternal and paternal survival, respectively, per year of schooling induced by the reform in the pooled sample (Table 2.3). All models are 2SLS regression models including the following controls: indicators for five-year age group, an indicator for census year (in the regressions for Mother Alive), indicator for heap year, indicators for province of birth, indicator for sex, and the interactions of sex with each of the other covariates. Exposure to the reform was used as an instrument for years of schooling. The instruments were an indicator for partial exposure to the reform (year of birth 1960 - 1966) and an indicator for full exposure to the reform (year of birth >= 1967). Model 1 includes a linear and quadratic term in age instead of the indicators for five-year age group. Model 4 includes cohorts born +/- 5 years of the reform cutoff. In model 6, the instrument was the interaction between the two reform indicators and a continuous variable for the proportion of respondents with no schooling by province of birth in the pre-reform period, while controlling for the main effects of each variable. Robust standard errors in parentheses. *F*-statistics were 18,780 and 402 (model 1); 14,923 and 52 (model 2); 15,602 and 69 (model 3); 125 and 62 (model 4); 33,611 (model 5); and 3,371 and 1,737 (model 6). *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Tanzania, born in Tanzania after 1945, and ages at least 15 years old. No weights were used. Source: Tanzania Census 1988 and 2002.

2.5.5 Spousal Survival and Survival Rates

Figure 2.2 displays the coefficients of schooling from 2SLS models assessing the causal effect of additional children's schooling on spousal survival among mothers co-residing with their children. The positive effects on survival primarily came from old age, and peaked for mothers ages 65 - 69 (3.8 percentage points, $p < 0.001$), net of time trends and other covariates. Table B6 in the supplementary materials shows results for the association between parental exposure to the reform and 14-year survival rates, and finds that additional exposure to the reform increases a parental cohort's probability of surviving by at least 4 percentage points ($p < 0.001$).

Fig. 2.2: *Coefficients of Children's Education in 2SLS Models for Spousal Survival*



Notes: Figure shows coefficients of schooling from 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. Coefficients represent the causal effect of one year of additional schooling on the old-age survival of spouses of mothers (and likely fathers) co-residing with their children in the 1988 and 2002 census. The instruments were an indicator for partial exposure to the policy reform (year of birth 1960 - 1966) and full exposure to the policy reform (year of birth ≥ 1967). All regressions included the following controls: indicators for five-year age groups, an indicator for census year, indicator for heap year, indicators for province of birth, indicator for sex, and the interactions of each covariate with sex. Separate 2SLS models were run for each of the maternal age categories. Although I focus on husbands whose wife is alive at the time of the census, this is the case for most Tanzanian men (97% in the 1988 census and 98% in the 2002 census). Dashed lines represent 95% confidence intervals. Sample includes survey respondents who were citizens of Tanzania, born in Tanzania after 1945, and ages at least 15 years old. Source: Tanzania Census 1988 and 2002.

2.5.6 Pathways

Table B8 in the supplementary materials shows 2SLS results for the effect of additional years of schooling on my secondary outcomes including: labor force participation, literacy, marital status, spousal years of schooling, household size, and household access to basic utilities, for women and men separately. Men saw a large increase in labor force participation (2.1 percentage points, $p < 0.001$); women somewhat less (1.5 percentage points, $p < 0.001$). Controlling for this pathway, however, in my main regressions did not substantially decrease the effect of additional years of schooling on parental survival (Table 2.7), suggesting that additional pathways may be likely (Table B8 in the supplementary materials). Adult literacy, for instance, increased by 9.9 percentage points in the pooled sample as a result of one year of additional schooling ($p < 0.001$).

Table 2.5: *OLS Results: Mediator Analysis of the Effect of Schooling on Parental Survival Including Labor Force Participation, Literacy, Marital Status, Spousal Years of Schooling, and Household Utilities*

Pathway	**Labor Force Participation (1-yes, 0=no)**	**Literacy (1=yes, 0=no)**	**Married (1=yes, 0=no)**	**Spousal Schooling (Years)**	**Household Utilities (1=yes, 0=no)**
Exposure	**Schooling**	**Schooling**	**Schooling**	**Schooling**	**Schooling**
Dependent Variable					
Mother Alive	0.0096***	0.0087***	0.0086***	0.0079***	0.0067***
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0016)
Father Alive	0.0092***	0.0100***	0.0099***	0.0098***	0.0067***
	(0.0001)	(0.0002)	(0.0002)	(0.0003)	(0.0017)
Additional Controls					
Labor Force Participation	✓	✓	✓	✓	✓
Literacy Child	-	✓	✓	✓	✓
Marital Status	-	-	✓	✓	✓
Spousal Schooling	-	-	-	✓	✓
Household Utilities	-	-	-	-	✓

Notes: Table shows regression results from OLS models controlling for indicators for five-year age groups, an indicator for census year (in the regressions for Mother Alive), indicator for heap year, indicators for province of birth, indicator for sex, and the interactions of each covariate with sex. Household utilities include access to a toilet or latrine, piped (running) water, and electricity. The sample for model 4 includes respondents living with their spouse. The sample for model 5 includes respondents living with their spouse and at least one parent. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights were used. Sample includes survey respondents who were citizens of Tanzania, born in Tanzania after 1945, and ages at least 15 years old. Source: Tanzania Census 1988 and 2002.

2.6 Discussion

The results presented in this paper suggest that children's educational attainment has a sizeable protective effect on their parents' old age survival. In the Tanzanian setting studied, each additional year of primary schooling in children resulted in an 11.7% reduction in the probability of maternal death, and a 10.4% reduction of paternal death by the time exposed child cohorts reach age 40, net of time trends and other covariates. I find that the survival benefits are larger for boys compared to girls, likely reflecting larger labor market returns to education for this subgroup. Estimated mortality reductions were consistent across a wide array of robustness checks. The key assumption underlying my empirical model is that the additional education attained by reform cohorts is orthogonal to other predictors of mortality conditional on time trends and covariates. I use a large number of tests to verify this, and apply a range of different time trends in my multivariable models, which should fully absorb gradual changes in children's characteristics over time. The main logic underlying my identification strategy is that the reform led to a large and unanticipated *additional* increase in human capital, which benefitted some cohorts more than others.

There are many reasons why parental survival might change across children's birth cohorts but the likely candidates—changes in socio-economic context, changes in public health prevention programming—are phenomena that are either gradual over time and/or affect children of many different ages: in both cases, these phenomena would result in gradual changes in parental survival across ages and would be picked up in my controls for underlying trends in age. To generate added confidence that my results are not confounded by other changes, I used an alternative identification strategy. I estimated difference-in-differences models exploiting geographical variation in impact of the reform. Indeed, the reform had a larger effect for children in birth provinces where a higher share of the population had no schooling in the pre-reform period.

While the educational reform allows me to identify the causal effects of interest, relying on changes generated by the reform comes with some limitations in external validity. First, the variation explored reflects mostly variation in primary schooling. Effects of schooling may be qualitatively and quantitatively different in secondary school and beyond, and indeed, in my descriptive analysis, I found the association between schooling and parental survival to be non-linear (Figure B3 in the supplementary materials). Second, the causal effects I identify do not necessarily represent the causal effects in the general population, but rather represent the causal effects among the subpopulation of compliers, i.e. children who attained increased schooling because of the reform. This subpopulation consists of children who, in the absence of the reform, would not have enrolled in primary school—a group more at risk of parental mortality compared to those who would have completed primary school regardless of the reform. Nevertheless, one advantage of using Tanzania's supply-side school reform is the important size of its impact (more than one year of additional schooling on a population level), and that it affected a large segment of the population who would otherwise not have enrolled in primary school.

2.7 Conclusion

These findings suggest that additional primary schooling in children does not only have large benefits for the children directly affected by it, but is also beneficial for other members of the core and potentially also extended family. Studies focusing on the private returns to schooling only are likely to severely underestimate the societal benefits of education, particularly in settings where welfare sectors are underdeveloped and within-family transfers are common.

Chapter 3

Causal Effect of Parental Schooling on Early Childhood Undernutrition — Quasi-experimental Evidence from Zimbabwe[1]

3.1 Introduction

Among children younger than 5 years, undernutrition is the leading risk factor worldwide, accounting for 12.4% of global disability-adjusted life years in 2010 (Lim et al. 2012). Previous cross-sectional and longitudinal studies have generally found a protective association between additional formal parental education and their children's nutritional status (Vollmer S et al. 2016, Phiri 2014, Silvestrin et al. 2013, Jacques Be-Ofuriyua 2011, Finlay, Ozaltin, and Canning 2011, Egata, Berhane, and Worku 2014, Corsi, Mejía-Guevara, and Subramanian 2015, Burchi 2012, Boyle et al. 2006, Behrman and Wolfe 1987), health (Behrman and Rosenzweig 2002, Behrman and Wolfe 1987, Finlay, Ozaltin, and Canning 2011, Victoria et al. 1992), and mortality risk (Cochrane, Leslie, and O'Hara 1982, Gakidou et al. 2010, Victoria et al. 1992). Hypothesized pathways include direct knowledge transfers to future parents (Aslam and Kingdon 2012, Glewwe 1999); increased ability to acquire, process (Christiaensen and Alderman 2004, Glewwe 1999), and communicate (Vikram, Vanneman, and Desai 2012) information; reduced risky health behaviors during pregnancy for mothers (e.g., smoking) (Currie and Moretti 2003) increased labor force participation and additional resources that could be invested in e.g., healthcare services (Bicego and Boerma 1993, Case, Lubotsky, and Paxson 2002, Cleland and van Ginneken 1988); assortative mating

[1] Co-authored with Subu Subramanian. This essay is forthcoming in part in the *American Journal of Epidemiology (De Neve and Subramanian 2017).*

with better educated and higher income spouses (Behrman and Rosenzweig 2004); a general familiarity with "modern society" (Glewwe 1999) that may make parents more receptive to modern medicine (Aslam and Kingdon 2012, Frost, Forste, and Haas 2005); and the "hidden curriculum" values of discipline and obedience of authority learned in school (Basu and Stephenson 2005). One key empirical concern with these previous studies is that a range of unobserved characteristics shared between parents and children, such as genetic traits or socioeconomic background, may make observational studies subject to confounding bias. Unobserved factors associated with higher economic or social prospects for a parent will both lead to higher parental education and improved child outcomes, making these earlier findings difficult to interpret.

In the absence of randomized trials assessing formal education (as opposed to e.g., nutrition education (Penny et al. 2005, Bhutta et al. 2008, Vazir et al. 2013)), a handful of studies have exploited natural experiments to address concerns about unobserved heterogeneity (Grossman 2006). In developed countries, additional college education for parents led to improved health among their children (Currie and Moretti 2003), while primary education for parents did not (Lindeboom, Llena-Nozal, and van der Klaauw 2009, McCrary and Royer 2011). Conceptually, large causal effects of parental education on child health also seem more likely in a developing setting, without an established welfare system generously providing support to households most at risk of early childhood undernutrition. A recent study in Turkey, exploiting a change in compulsory schooling law as a natural experiment, found that maternal primary school completion increased height-for-age Z-scores (HAZ) and weight-for-age Z-scores (WAZ) of their children by 1.1 and 1.0 standard deviations, respectively. Children born to mothers with primary school education were 2 inches taller on average relative to children born to mothers without primary school education (Güneş 2015). Evidence on the causal relationship between additional education for both parents and undernutrition in children in low-resource settings, however, remains scarce. This study is among the first to estimate a causal effect of formal parental schooling on anthropometric failure in children in a low-resource setting.

In this study, I use data from the Demographic and Health Surveys (DHS) from Zimbabwe in a first attempt to address this gap in the literature. To overcome the confounding concerns outlined above, I exploit quasi-experimental variation in educational attainment generated by one of the largest education policy reforms in low- and middle-income settings. In 1980, Zimbabwe eliminated barriers to secondary school which affected all parental cohorts born in 1967 or later, but not parents born earlier. As shown in this essay, the reform increased educational attainment by an average of 2.5 years for mothers and 2.4 years for fathers fully affected by the reform. Since there is no obvious reason why parents of children born just before 1967 would be different from parents of children born just after, conditional on covariates (such as generic time trends in childhood nutrition risk across parental birth cohorts), I can directly test the anthropometric impact of exogenous shifts in parental educational attainment on children's outcomes.

3.2 Methods

3.2.1 Data Sources

Data were obtained from the 1988, 1999, 2005-06, and 2010-11 DHS surveys, four nationally-representative household surveys with anthropometric data for children ages under 5 at the time of the survey. For each survey, between 4,000 and 10,000 households were selected; all members aged 15 – 49 for women and all members aged 15 – 54 for men were eligible to be interviewed. Household response rates ranged from 95% to 99%; individual participation ranged from 90% to 96% (women) and from 82% to 92% (men), yielding a total sample of 146,630 individuals (Demographic and Health Surveys 2011). Children were eligible for anthropometric measurements if they were between 3 and 59 months old and living in the household at the time of the survey. Data on parental age and completed years of schooling were available for 88% of children eligible for anthropometric measurement.

3.2.2 Study Population

Zimbabwe has among the highest rates of child undernutrition in the world, with 32% stunted, 10% underweight, and 3% of children under 5 wasted in 2011; in addition, overweight is increasingly becoming a concern, with 6% of children under 5 overweight, and 31% of women and 9% of men ages 15 - 49 either overweight (BMI ≥ 25 m/kg^2) or obese (BMI ≥ 30 m/kg^2) (Demographic and Health Surveys 2011). The prevalence of "double burden" malnutrition in households is relatively uncommon (e.g., under 2% for obese mothers and underweight, stunted or wasted child pairs in 2005 (Wojcicki 2014)) (see Supplementary Appendix for additional details on nutrition context). The study population included all children ages between 3 and 59 months residing in the country and living with at least one parent at the time of the survey. I excluded children with parents born before 1957 and after 1977 to maximize the comparability of pre- and post-reform cohorts. Figure C1 in the Supplementary Appendix displays the study participant flow diagram.

3.2.3 Outcomes

The outcome of interest was anthropometric failure of the child at the time of the survey, defined as early childhood stunting (i.e., too short for age, an indicator of chronic malnutrition), wasting (i.e., too thin for height, an indicator of acute malnutrition), and underweight (i.e., weight-for-age, a composite index). As a baseline, I show anthropometric measures based on the U.S. National Center for Health Statistics (NCHS) growth reference, since most DHS surveys in Zimbabwe were conducted before the introduction of new growth standards in 2006 by the World Health Organization (WHO) (Borghi et al. 2006, Group 2006). Anthropometric failure was defined as 2 or more standard deviations below the median of an international reference population defined by the NCHS. A child below minus two standard deviations (-2 SD) from the median of the NCHS reference population in terms of height-for-age was considered stunted. Similar calculations were done for wasting and underweight. Biologically implausible values (defined by the NCHS for height as a Z score of less than –6 or greater than 6, for weight as a Z score of less than –6 or greater than 6, and for weight-for-height as a Z score of less than –4 or greater than 6) were excluded (Borghi et al. 2006). I show a wide range of additional outcomes to generate added confidence in my findings. First, I tested the sensitivity of my results to the new WHO growth standards (Borghi et al. 2006), using separate files released by the DHS which contain the new z-scores for older DHS surveys. Second, I assessed continuous measures for HAZ, WAZ, and weight-for-height Z-scores, "severe" undernutrition (defined as Z scores of less than –3), and overweight (defined as Z scores of more than 2).

3.2.4 Exposure

The key exposure in my analysis was total parental "years of schooling completed" by the time of the survey. In all models, I assessed maternal and paternal schooling separately.

3.2.5 Exogenous Instrument

To obtain causal effects, I exploited exogenous variation in educational attainment resulting from a 1980 policy reform that eliminated restrictions limiting the advancement toward secondary school. In 1980, following independence, Zimbabwe rapidly expanded access to secondary schools for black Zimbabweans (see Supplementary Appendix for background information on the education system and reform). This 'natural experiment' provides an opportunity to estimate the causal impact of parental schooling on early childhood undernutrition, by comparing children of parental birth cohorts exposed to the reform vs. those unexposed. I defined an indicator–"fully exposed"–taking the value one if the parent was born in a cohort exposed to the 1980 education policy reform and zero otherwise (YOB ≥ 1967); and an indicator–"partially exposed"–if the parent was partially exposed to the reform (1963 < YOB < 1967). Given that children are expected to start primary school at age 7, they are expected to enter lower secondary school at age 13. Therefore, parents born in 1967 or later would have entered secondary school in 1980 or later, and were thus classified as fully exposed.

3.2.6 Covariates

I controlled flexibly for parental age with single-year age indicators to account for the lower expected years of schooling at younger ages. I also included a continuous linear term in parental year of birth to account for temporal trends across birth cohorts. Critically, the availability of multiple survey rounds allowed me to generate variation in parental age in a given parental birth cohort group. I implicitly also control for period effects by simultaneously adjusting for parental age and year of birth. In sensitivity analyses, I added indicators for child characteristics (age, sex, and birth order) and region as additional controls in my main model.

3.2.7 Statistical Analyses

My analysis of the education reform proceeded in three steps. First, I assessed whether parental birth cohorts exposed to the reform had higher educational attainment than birth cohorts not exposed to the reform. Second, I assessed the ITT effect of being in a parental reform cohort on undernutrition in their children. Third, I estimated 2SLS regression models, using parental exposure to the reform as an instrument for years of schooling while adjusting for covariates. Natural experiments that change the probability of an exposure can be analyzed like RCT's with non-compliance (Little and Rubin 2000). Under plausible assumptions, the treatment effect among compliers is the ratio of the ITT and the difference in the probability of receiving treatment. These IV estimates are interpreted as specific to the subpopulation who complied with their treatment assignment – i.e. parents who increased their years of schooling because of the policy reform (Angrist, Imbens, and Rubin 1996). As a benchmark for my 2SLS estimates, I assessed the naïve association between parental schooling and undernutrition in their children. I assessed the naïve relationship graphically and then adjusted for covariates.

3.2.8 Assumptions for Causal Inference

Four assumptions underpin causal inference in my study (see Figure C2 in the Supplementary Appendix for a directed acyclic graph). First, the policy change (instrument) must have an effect on educational attainment. Indeed, I found a large effect of the policy reform on total years of schooling completed among parents in my study. Second, parental birth cohorts just before and after the reform cut-off must be similar, conditional on covariates such as temporal trends. I conduct a wide range of robustness checks to generate added confidence in this assumption (described below). Third, I assume that the policy reform affected early childhood undernutrition only through parental schooling – this assumption is highly plausible since the reform was a supply-side intervention that is unlikely to have differentially affected the same birth cohorts, except through their increased access to school. Fourth, I assume monotonicity – i.e., exposure to the reform only induced parents to obtain more schooling or to have no change in schooling (Angrist, Imbens, and Rubin 1996, Hernan and Robins 2006).

3.2.9 Robustness Checks

As a first robustness check, I reduced the window of observation to a narrower set of parental birth cohorts (3-year and 5-year birth cohort windows around the reform cutoff). Underlying trends are more likely to be approximately linear when a narrower window of birth cohorts is included in the analytical sample. I assessed the robustness of my results to the presence of non-linearities in long-run cohort trends in risk of childhood undernutrition, by controlling for a quadratic in year of birth in addition to the linear term included in the main analysis. I used an alternative identification strategy using differential impact of the policy reform by geographical region. I also controlled for an indicator for heap year to take into account measurement error in parental age due to age heaping. I excluded ages that are multiples of five, since I still had all parental years of birth in my analysis sample (the surveys were conducted in more than five different years). I assessed the robustness of my results to the exclusion of parental cohorts that were partially exposed to the reform, including only parents with at least primary schooling (i.e., those most likely to have benefitted from the reform), including the full sample of parental cohorts, modeling the outcome using a Probit link function, and using sampling weights (Deaton 1997). Finally, I implemented a placebo test. I shifted the cutoff for the treatment and control cohorts by 5 years forward. With this placebo test, the first stage should become insignificant and there should be no or vastly decreased difference between the education levels of young and old parental cohorts.

3.2.10 Heterogeneity

I assessed heterogeneity of the effect of parental schooling on undernutrition in their children. I ran my main model in a range of subsamples, including in households with a high (quintiles 4 – 5) and low (quintiles 1 – 2) wealth index, in an urban and rural setting, and restricting the sample of children to alternative age groups (3 - 36 months, 3 - 24 months).

This study was exempted from full review by the Harvard T.H. Chan School of Public Health Institutional Review Board since it uses de-identified data. All analyses were conducted in Stata 14 (StataCorp, College Station, Texas, USA).

3.3 Results

The 1988, 1999, 2005-06, and 2010-11 DHS surveys included 8,243 children with data on anthropometric failure and parental characteristics; 6,428 with an eligible mother and 4,175 with an eligible father (Table 3.1, Figure C1).

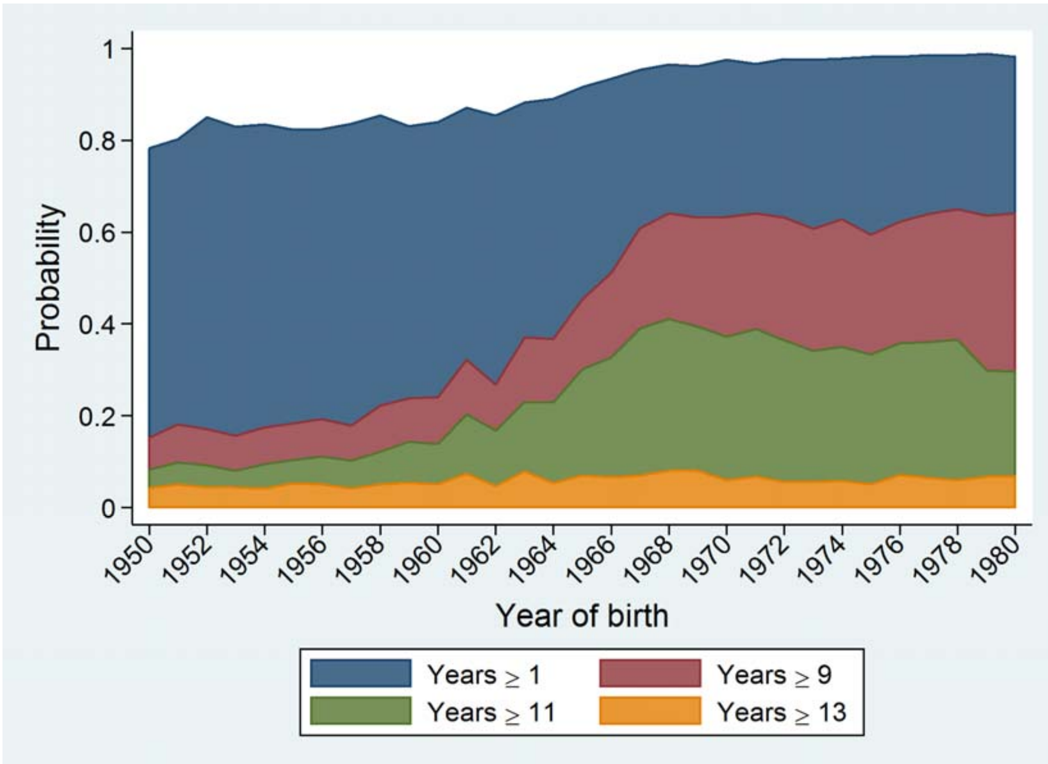
Table 3.1: *Selected Characteristics of the Study Participants*

Subsample	**Pre-reform cohorts (1957 ≤ parental YOB ≤ 1962)**	**Partially affected (1963 ≤ parental YOB ≤ 1966)**	**Post-reform cohorts (1967 ≤ parental YOB ≤ 1977)**
Children 3 - 59 months old			
		With eligible mother (N=6,428)	
Age — mo	32.7±16.6	31.4±16.6	30.6±16.5
Female — no. (%)	604 (49.6)	562 (50.9)	2,053 (50.6)
Stunted — no. (%)	342 (28.6)	294 (27.3)	1,163 (28.6)
Underweight — no. (%)	149 (11.8)	120 (10.7)	638 (15.2)
Wasted — no. (%)	42 (3.5)	30 (2.6)	203 (4.9)
		With eligible father (N=4,175)	
Age — mo	33.5±16.7	32.7±16.7	30.0±16.6
Female — no. (%)	327 (50.3)	302 (52.6)	1,475 (50.3)
Stunted — no. (%)	175 (27.9)	152 (26.8)	791 (26.8)
Underweight — no. (%)	89 (14.1)	85 (14.6)	425 (14.0)
Wasted — no. (%)	37 (5.7)	30 (4.9)	149 (4.8)
Parents			
		Mothers (N=9,164)	
Age — years	37.0±7.3	32.6±7.1	28.9±6.1
At least 1 year of schooling — no. (%)	1,498 (77.9)	1,346 (85.7)	5,391 (95.6)
Schooling — years	4.7±3.6	6.0±3.8	8.1±3.1
Urban status — no. (%)	413 (24.0)	400 (27.3)	1,535 (30.1)
Wealth index — quintile (1-5)	2.6±1.4	2.9±1.4	3.0±1.4
		Fathers (N=5,312)	
Age — years	40.4±6.2	36.6±6.1	32.4±5.0
At least 1 year of schooling — no. (%)	1,049 (93.2)	848 (97.2)	3,265 (99.0)
Schooling — years	7.5±3.7	9.1±3.5	9.6±2.8
Urban status — no. (%)	345 (34.0)	305 (38.9)	1,239 (40.9)
Wealth index — quintile (1-5)	3.0±1.5	3.1±1.5	3.2±1.5

Notes: Plus–minus values are means ± SD. YOB: year of birth. The sample for children includes children ages between 3 and 59 months at the time of the survey, living with a parent born 1957 - 1977, and with valid anthropometric data. The sample for parents includes parents born between 1957 and 1977 who lived with their child under age 5. Men were not included and wealth index was not available in the 1988 survey. Sample weights used as provided. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

The reform increased average years of schooling by 2.5 years ($p < 0.001$) for mothers and 2.4 years ($p < 0.001$) for fathers fully exposed to the reform, with the effect largely driven by gains in lower secondary school completion (Table C1 in the Supplementary Appendix). Figure 3.1 shows the percentage of respondents for each birth cohort who completed at least 1, 9, 11, and 13 years of schooling. The percentage of each cohort with at least 9 or 11 years of schooling increased dramatically for cohorts born in 1967 or later. Due to grade repetition and/or late entry into school, some respondents born between 1963 and 1966 likely were also affected by the reform.

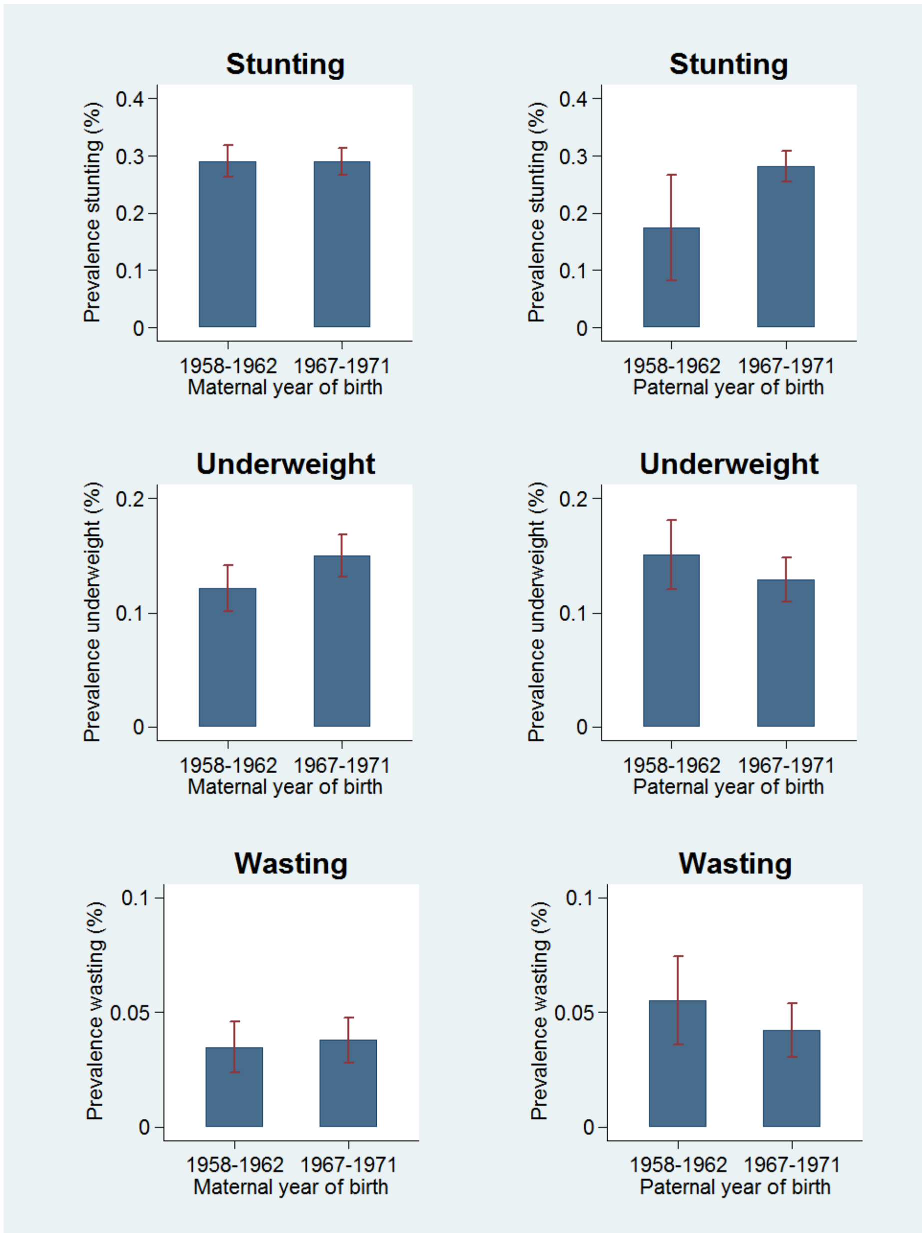
Fig. 3.1: *Adult Educational Attainment by Birth Cohort in Zimbabwe*



Notes: Figure shows the probability that the respondent has completed at least X years of formal schooling. Schooling was defined as "total years of schooling completed" by the time of the survey. Sample includes survey respondents at least 21 years old at the time of the surveys. Survey weights used as provided. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

Table C2 in the Supplementary Appendix presents ITT results, in which childhood undernutrition was regressed directly on the instrument and covariates. I find no significant differences in the probability of early childhood stunting, wasting, and underweight for children of parents that were exposed to the reform compared to children of parents that were not exposed. Figure 2 displays undernutrition by parental year of birth and exposure to the education reform, separately for maternal and paternal exposure.

Fig. 3.2: *Childhood Undernutrition by Parental Birth Cohort and Exposure to the Reform*



Notes: Figure displays undernutrition by parental year of birth and exposure to the reform, separately for maternal and paternal exposure. The exposure of interest is whether the parent was born in 1958-1962 (pre-reform) or in 1967-1971 (post-reform). Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. Survey weights used as provided. Error bars represent 95% confidence intervals. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

The crude association between additional parental schooling and undernutrition in their children was non-linear, in particular among fathers: children of fathers with 6 to 7 years of education had the highest risk, declining sharply thereafter for children of fathers with 8 years of education or more (Figure C3 in the Supplementary Appendix). The association between parental schooling and anthropometric failure help up in multivariable OLS regression models (Model 1, Table 3.2). Each additional year of maternal schooling was associated with a decrease of 1.6 percentage points ($p < 0.001$) in the probability of her child being stunted, 1.1 percentage points ($p < 0.001$) in being underweight, and 0.3 percentage points in being wasted ($p < 0.001$). Each additional year of paternal schooling was associated with a decrease of 1.0 percentage points in the probability of his child being stunted ($p < 0.001$), 0.9 percentage points in being underweight ($p < 0.001$), and 0.2 percentage points in being wasted ($p = 0.07$). In contrast to the negative association between parental schooling and all three measures of anthropometric failure in OLS regressions, I find no evidence of a causal effect of parental schooling in 2SLS (IV) regressions (Model 2, Table 3.2).

Table 3.2: *OLS and 2SLS Regression Results: Parental Schooling and Childhood Undernutrition*

Exposure	**Maternal schooling**			**Paternal schooling**		
Outcome	**Stunted**	**Underweight**	**Wasted**	**Stunted**	**Underweight**	**Wasted**
Model 1: OLS						
Effect estimate:	-0.016***	-0.011***	-0.003***	-0.010***	-0.009***	-0.002*
years of schooling (SE)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Model 2: 2SLS (IV)						
Effect estimate:	0.015	0.016	-0.005	0.017	0.005	-0.007
years of schooling (SE)	(0.014)	(0.011)	(0.007)	(0.017)	(0.013)	(0.008)
No. of children	6,428	6,428	6,428	4,175	4,175	4,175

Notes: Model 1 is an ordinary least squares (OLS) linear probability model. Model 2 is a two-stage least squares (2SLS) linear probability model, in which exposure to the reform was used as an instrumental variable (IV) for parental years of schooling. All models included the following controls: an indicator for parental age, a linear trend in parental year of birth (YOB), and an indicator for survey year. The instruments were an indicator for partial exposure to the reform (parental YOB 1963 - 1966) and an indicator for full exposure to the reform (parental YOB >= 1967). Robust standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes children ages between 3 and 59 months at the time of the survey and living with a parent born 1957 - 1977. *F*-statistics were 49.6 for maternal schooling and 33.3 for paternal schooling. No weights were used. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

3.3.1 Sensitivity Analyses

Table 3.3 and Tables C3 to C7 in the Supplementary Appendix display the results of sensitivity analyses for the ITT and 2SLS results. In general, results were not sensitive to accounting for parental age heaping, different modeling strategies for parental year of birth, different specifications and definitions of the outcome, alternate sample specifications (a narrower birth cohort window around the reform cut-off, excluding parents that were partially exposed to the reform, those with at least seven years of schooling, and the full sample of parents), the inclusion of additional controls for a child's age, sex, and birth order, region indicators, using sampling weights, or using an alternative identification strategy (Tables 3.3 and C3 in the Supplementary Appendix). A small effect was seen for additional maternal schooling on the probability of her child being underweight and wasted using alternate sample specifications, on the probability of her child being underweight using the new WHO criteria, and for additional paternal schooling on WAZ and WHZ, but these coefficients did not reach conventional benchmarks for statistical significance in any of the alternative specifications. Using the new WHO criteria, additional maternal schooling slightly increased the probability of her child being stunted (3.5 percentage points, $p = 0.02$) (Table C5 in the Supplementary Appendix). In a placebo test, using an arbitrary cutoff for the education reform, I found that the first stage became insignificant and that there was no difference in educational attainment between the pre- and post-reform cohorts (Table C1 in the Supplementary Appendix). Similar to my main results, the overall impression of these estimates is that parental schooling had no effect on early childhood undernutrition at the national level.

Table 3.3: *Sensitivity Analyses Regression Results: Assessing the Robustness of my Results to Indicators for Region, Year of Birth Squared, Sample Weights, Probit, Alternative Samples and Identification Strategy*

Sensitivity analysis	**Region**	**Quadratic year of birth**	**Sample weights**	**Probit**	**5-year cohort window**	**Excluding partially exposed**	**Full sample**	**Diff-in-diff**
Exposure	**Reform indicator (SE)**	**Reform indicator (SE)**	**Reform indicator (SE)**	**Reform indicator (SE)**	**Reform indicator (SE)**	**Reform indicator (SE)**	**Reform indicator (SE)**	**Years of schooling (SE)**
				Maternal exposure				
Outcome								
Stunted	0.040	0.064	0.033	0.043	0.144	0.045	0.012	0.012
	(0.038)	(0.058)	(0.043)	(0.038)	(0.123)	(0.041)	(0.021)	(0.029)
Underweight	0.036	0.032	0.022	0.034	-0.015	0.033	0.015	0.009
	(0.030)	(0.043)	(0.033)	(0.029)	(0.082)	(0.032)	(0.016)	(0.022)
Wasted	-0.020	-0.015	-0.029	-0.022	0.001	-0.035*	-0.016*	0.017
	(0.018)	(0.027)	(0.020)	(0.017)	(0.051)	(0.020)	(0.009)	(0.013)
No. of children	6,428	6,428	6,428	6,428	2,799	5,469	13,461	6,428
				Paternal exposure				
Outcome								
Stunted	0.056	0.074	0.070	0.050	0.158	0.046	0.019	-0.016
	(0.042)	(0.069)	(0.046)	(0.041)	(0.128)	(0.042)	(0.031)	(0.035)
Underweight	0.005	-0.001	-0.006	-0.002	-0.054	-0.009	0.012	-0.019
	(0.032)	(0.052)	(0.035)	(0.032)	(0.089)	(0.033)	(0.024)	(0.027)
Wasted	-0.018	-0.007	-0.018	-0.025	-0.060	-0.022	0.000	-0.020
	(0.020)	(0.031)	(0.020)	(0.021)	(0.039)	(0.020)	(0.015)	(0.020)
No. of children	4,175	4,175	4,175	4,175	1,848	3,592	6,832	4,175

Notes: Table presents robustness checks for the results presented in Table 3.2. Models 1 - 7 are regression results from intention-to-treat (ITT) OLS models, controlling for an indicator for parental age, a linear trend in parental year of birth (YOB), an indicator for survey year, and an indicator for partial exposure to the reform (YOB 1963 - 1966). Full exposure to the reform was defined as 1 if parental YOB >= 1967, zero otherwise. Model 2 additionally controls for parental YOB squared. Model 4 is a Probit model. Model 8 is a two-stage least squares (2SLS) model, where the instrumental variable (IV) was the interaction between a reform indicator and a continuous variable for the proportion of respondents with no schooling by region in the pre-reform period, while controlling for the main effects of each variable. Robust standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No sample weights were used, except in model 3. Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

3.3.2 Heterogeneity

Table 3.4 shows 2SLS results for the effect of parental schooling on undernutrition in their children in subsamples. Results were consistent across children in households that were most at risk of early childhood undernutrition: those living in rural settings, in the lowest wealth quintiles (1 – 2), and children aged less than 36 months (Figure C4 in the Supplementary Appendix). For households in the urban and high wealth index subsamples, each additional year of maternal schooling induced by the policy reform led to an absolute reduction in the probability of her child being wasted of 4.8 percentage points (p = 0.01) and 3.2 percentage points (p = 0.03), respectively; for children ages between 3 and 23 months, each additional year of paternal schooling increased the probability of his child being stunted by 7.2 percentage points (p = 0.04).

Table 3.4: *2SLS Regression Results in Subsamples: Parental Schooling and Undernutrition in their Children*

Exposure	**Maternal schooling**			**Paternal schooling**		
Outcome	**Stunted**	**Underweight**	**Wasted**	**Stunted**	**Underweight**	**Wasted**
Model: 2SLS (IV)						
			Urban subsample (N=2,108)			
Effect estimate:	0.028	0.032	-0.048***	0.021	-0.017	0.003
years of schooling (SE)	(0.037)	(0.030)	(0.019)	(0.032)	(0.021)	(0.012)
			Rural subsample (N=6,141)			
Effect estimate:	0.010	0.016	0.006	0.023	0.018	-0.007
years of schooling (SE)	(0.016)	(0.012)	(0.008)	(0.021)	(0.017)	(0.011)
			High wealth index subsample (N=2,226)			
Effect estimate:	-0.013	0.011	-0.032**	0.013	-0.027	0.001
years of schooling (SE)	(0.028)	(0.022)	(0.015)	(0.037)	(0.025)	(0.015)
			Low wealth index subsample (N=3,250)			
Effect estimate:	0.003	0.017	0.000	0.020	0.023	0.002
years of schooling (SE)	(0.021)	(0.018)	(0.011)	(0.020)	(0.017)	(0.010)
			Children 3 - 35 months subsample (N=4,810)			
Effect estimate:	0.003	0.004	-0.010	0.014	-0.010	-0.011
years of schooling (SE)	(0.021)	(0.017)	(0.011)	(0.020)	(0.016)	(0.011)
			Children 3 - 23 months subsample (N=3,054)			
Effect estimate:	0.001	0.004	-0.003	0.072**	0.005	-0.014
years of schooling (SE)	(0.027)	(0.021)	(0.015)	(0.035)	(0.024)	(0.018)

Notes: Regression results from a two-stage least squares (2SLS) linear probability model, in which exposure to the reform was used as an instrumental variable (IV) for parental years of schooling. The model included the following controls: an indicator for parental age, a linear trend in parental year of birth (YOB), and an indicator for survey year. The instruments were an indicator for partial exposure to the reform (parental YOB 1963 - 1966) and an indicator for full exposure to the reform (parental YOB >= 1967). Low Wealth Index was defined as a household wealth index in quintiles 1 and 2 and High Wealth Index as a household wealth index in quintiles 4 and 5. Wealth index was not available in the 1988 survey. Robust standard errors (SE) in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. *F*-statistics for maternal and paternal schooling, respectively, were 7.7 and 10.8 (urban subsample), 45.7 and 22.9 (rural subsample), 11.7 and 7.3 (high wealth index), 25.5 and 28.7 (low wealth index), 18.2 and 21.4 (children 3 – 35 months), and 12.2 and 9.2 (children 3 – 23 months). No weights were used. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11

3.4 Discussion

Using an education policy reform as a natural experiment, I find that additional parental schooling has no effect on the probability of stunting, wasting, and underweight in their children under 5 at the national level in Zimbabwe. Findings were consistent across a wide range of sensitivity analyses. For households in the urban and high wealth index subsamples, each additional year of maternal schooling had a protective effect against the probability of being wasted in children; whereas among children ages 3 through 23 months, each additional year of paternal schooling increased the probability of being stunted. I interpret my 2SLS (IV) estimates as causal because they are not vulnerable to unobserved factors – e.g. genetic traits, socio-economic background – that may confound previous studies of the association between parental education and early childhood undernutrition (Vollmer et al. 2016, Phiri 2014, Silvestrin et al. 2013, Jacques Be-Ofuriyua 2011, Finlay, Ozaltin, and Canning 2011, Egata, Berhane, and Worku 2014, Corsi, Mejía-Guevara, and Subramanian 2015, Burchi 2012, Boyle et al. 2006, Behrman and Wolfe 1987). The main logic underlying my identification strategy is that the reform led to a large and unanticipated additional increase in parental schooling, which benefitted some birth cohorts more than others. Critically, the availability of multiple survey rounds allowed me to include controls for age effects and continuous temporal trends (such as gradual changes in the nutritional context). I also used an alternative identification strategy which yielded similar results, and conducted a placebo test comparing cohorts on either side of an arbitrary reform and found no difference between the education levels of young and old parental cohorts.

This study has a few limitations with regard to external validity. First, most of the variation explored reflects variation in secondary schooling. Effects of schooling may be qualitatively and quantitatively different in primary school (Makoka and

Masibo 2015). In my descriptive analysis, I found the association between parental schooling and childhood undernutrition to be non-linear, particularly among fathers: an additional year of primary school was associated with an increase in the probability of stunting and underweight, but secondary school and beyond was associated with a reduction in the probability of stunting and underweight (Figure C3 in the Supplementary Appendix). Second, the causal effects that I estimate are specific to the subpopulation of "compliers" – i.e. those induced to increase schooling because of the reform. This subpopulation consists of parents who, in the absence of the reform, would have dropped out after seventh grade. Third, the results are specific to the context and years of exposure under study (the 1970s through 1990s). One advantage of using the education reform was the important size of its impact and that it affected a large proportion of the population. These findings are relevant to many other countries in sub-Saharan Africa, where a large number of similar reforms have dramatically increased access to school.

The effects of parental schooling on early childhood undernutrition have been observed in other settings. In a Turkish study, maternal primary school completion was associated with a 0.2 increase in HAZ and WAZ in their children in multivariable OLS models. This corresponds to the regression results from my OLS models. To obtain causal estimates, the study exploited a change in compulsory schooling law as a natural experiment but, in contrast to the IV estimates presented here, found sizeable positive effects of maternal primary school completion on HAZ and WAZ—an increase of about one standard deviation (Güneş 2015). One reason for these positive results could be that the margin of schooling affected by this compulsory schooling law was a key margin for spillover effects to children. In another natural experiment from Zimbabwe, each additional year of maternal schooling decreased child mortality by about 21%, but this effect was primarily driven by changes in age at first cohabitation and sexual intercourse (Grépin and Bharadwaj 2015). Although increased secondary school among mothers affected a number of pathways, including sexual debut, age at first birth, marriage, and economic opportunities, the study found little evidence supporting other channels through which maternal schooling might affect child health (such as health seeking behavior).

Schooling may not impact all subpopulations equally for several reasons. Effects may depend, for instance, on the degree of foresight in the subpopulation affected by the policy reform and the various pathways by which schooling affects child health (McCrary and Royer 2011). Parents might not invest in behaviors protective of child health if they are unable to anticipate the important labor market (Bennell and Malaba 1993) or non-pecuniary (Bor and De Neve 2015, De Neve et al. 2015) returns to schooling. Stunting and wasting also represent distinct biological processes and the pathways by which parental schooling may influence these outcomes are likely different. In further analysis, I found protective effects of maternal schooling on early childhood wasting in urban and wealthier households (but not on early childhood stunting or underweight), suggesting that in settings with increased resources and access to healthcare services, maternal schooling "can be" a determinant of early childhood wasting. Additional maternal schooling and access to resources might be complements in reducing the risk of childhood wasting, rather than substitutes as has been argued elsewhere (Rosenzweig and Schultz 1982). While I caution against generalizing to other contexts, my estimates are likely informative of the limits of providing schooling alone in a low-income setting, where childhood stunting in particular is endemic.

Expanding access to secondary school had no effect on childhood stunting, wasting, and underweight at the national level in Zimbabwe. My findings stand in contrast to what has long been suspected: parental schooling alone is not a 'panacea' against childhood undernutrition. This study further emphasizes the need for directed health interventions (Ruel and Alderman 2013) in low-resource settings.

References

Adamchak, D. J., A. O. Wilson, A. Nyanguru, and J. Hampson. 1991. "Elderly Support and Intergenerational Transfer in Zimbabwe: An Analysis by Gender, Marital Status, and Place of Residence." *The Gerontologist* 31 (4):505-513. doi: 10.1093/geront/31.4.505.

Agüero, J., and P. Bharadwaj. 2014. "Do the More Educated Know More about Health? Evidence from Schooling and HIV Knowledge in Zimbabwe." *Economic Development and Cultural Change* 62 (3).

Agwanda, A., and H Amani. 2014. "Population Growth, Structure and Momentum in Tanzania." *THDR 2014: Background Paper No. 7, ESRF Discussion Paper 61*.

Ainsworth, M., K. Beegle, and A. Nyamete. 1996. "The Impact of Women's Schooling on Fertility and Contraceptive Use: A Study of Fourteen Sub-Saharan African Countries." *The World Bank Economic Review* 10 (1):85-122. doi: 10.1093/wber/10.1.85.

Alsan, M. M., and D. M. Cutler. 2013. "Girls' education and HIV risk: evidence from Uganda." *Journal of Health Economics* 32 (5):863-72. doi: 10.1016/j.jhealeco.2013.06.002.

Angrist, J. D., G. W. Imbens, and D. B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434):444. doi: 10.2307/2291629.

Aslam, M., and G. G. Kingdon. 2012. "Parental Education and Child Health—Understanding the Pathways of Impact in Pakistan." *World Development* 40 (10):2014-2032. doi: 10.1016/j.worlddev.2012.05.007.

Baird, S. J., R. S. Garfein, C. T. McIntosh, and B. Ozler. 2012. "Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial." *The Lancet* 379 (9823):1320-9. doi: 10.1016/S0140-6736(11)61709-1.

Barnighausen, T., V. Hosegood, I. M. Timaeus, and M. L. Newell. 2007. "The socioeconomic determinants of HIV incidence: evidence from a longitudinal, population-based study in rural South Africa." *AIDS* 21 Suppl 7:S29-38. doi: 10.1097/01.aids.0000300533.59483.95.

Basu, A. M., and R. Stephenson. 2005. "Low levels of maternal education and the proximate determinants of childhood mortality: a little learning is not a dangerous thing." *Social Science & Medicine* 60 (9):2011-2023. doi: 10.1016/j.socscimed.2004.08.057.

Bauman, L. J., G. Foster, E. Johnson Silver, R. Berman, I. Gamble, and L. Muchaneta. 2006. "Children caring for their ill parents with HIV/AIDS." *Vulnerable Children and Youth Studies* 1 (1):56-70. doi: 10.1080/17450120600659077.

Becker, S. O., F. Cinnirella, and L. Woessmann. 2010. "The trade-off between fertility and education: evidence from before the demographic transition." *Journal of Economic Growth* 15 (3):177-204. doi: 10.1007/s10887-010-9054-x.

Behrman, J. R., and M. R. Rosenzweig. 2002. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review* 92 (1):323-334. doi: 10.1257/000282802760015757.

Behrman, J. R., and M. R. Rosenzweig. 2004. "Returns to Birthweight." *Review of Economics and Statistics* 86 (2):586-601. doi: 10.1162/003465304323031139.

Behrman, J. R., and B. L. Wolfe. 1987. "How does mother's schooling affect family health, nutrition, medical care usage, and household sanitation?" *Journal of Econometrics* 36 (1-2):185-204. doi: 10.1016/0304-4076(87)90049-2.

Bennell, P., and J. Malaba. 1993. "Up-dating rates of return to education: Zimbabwe in the late 1980s." *International Journal of Educational Development* 13 (3):277-287. doi: 10.1016/0738-0593(93)90035-x.

Berkman, L. F., T. Glass, I. Brissette, and T. E. Seeman. 2000. "From social integration to health: Durkheim in the new millennium." *Social Science & Medicine* 51 (6):843-57.

Berkman, L. F., and I. Kawachi. 2000. *Social epidemiology*. New York: Oxford University Press.

Bhutta, Z. A., T. Ahmed, R. E. Black, S. Cousens, K. Dewey, E. Giugliani, B. A. Haider, B. Kirkwood, S. S. Morris, H. P. S. Sachdev, and M. Shekar. 2008. "What works? Interventions for maternal and child undernutrition and survival." *The Lancet* 371 (9610):417-440. doi: 10.1016/s0140-6736(07)61693-6.

Bicego, G. T., and J. T. Boerma. 1993. "Maternal education and child survival: a comparative study of survey data from 17 countries." *Social Science & Medicine* 36 (9):1207-27.

Black, S. E., P. J. Devereux, and K. G. Salvanes. 2008. "Staying in the Classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births*." *The Economic Journal* 118 (530):1025-1054. doi: 10.1111/j.1468-0297.2008.02159.x.

Bloom, D. E., and D. Canning. 2004. "Global Demographic Change: Dimensions and Economic Significance." *National Bureau of Economic Research Working Paper Series* No. 10817. doi: 10.3386/w10817.

Bloom, D. E., D. Canning, and G. Fink. 2011. "Implications of population ageing for economic growth." *Oxford Review of Economic Policy* 26 (4):583-612. doi: 10.1093/oxrep/grq038.

Bloom, D. E., D. Canning, G. Fink, and J. E. Finlay. 2009. "Fertility, female labor force participation, and the demographic dividend." *Journal of Economic Growth* 14 (2):79-101. doi: 10.1007/s10887-009-9039-9.

Bor, J., and J. W. De Neve. 2015. "A Social Vaccine? HIV Infection, Fertility, and the Non-Pecuniary Returns to Secondary Schooling in Botswana." Population Association of America Annual Meeting, San Diego.

Bor, J., F. Tanser, M. L. Newell, and T. Barnighausen. 2012. "In A Study Of A Population Cohort In South Africa, HIV Patients On Antiretrovirals Had Nearly Full Recovery Of Employment." *Health Affairs* 31 (7):1459-1469. doi: 10.1377/hlthaff.2012.0407.

Borghi, E., M. de Onis, C. Garza, J. Van den Broeck, E. A. Frongillo, L. Grummer-Strawn, S. Van Buuren, H. Pan, L. Molinari, R. Martorell, A. W. Onyango, J. C. Martines, and W. H. O. Multicentre Growth Reference Study Group. 2006. "Construction of the World Health Organization child growth standards: selection of methods for attained growth curves." *Statistics in Medicine* 25 (2):247-65. doi: 10.1002/sim.2227.

Borkum, E. 2009. "Grade structure, educational attainment and labor market outcomes: Evidence from Botswana." *Job Market Paper, Columbia University*.

Boyle, M. H., Y. Racine, K. Georgiades, D. Snelling, S. Hong, W. Omariba, P. Hurley, and P. Rao-Melacini. 2006. "The influence of economic development level, household wealth and maternal education on child health in the developing world." *Social Science & Medicine* 63 (8):2242-2254. doi: 10.1016/j.socscimed.2006.04.034.

Bravo, J. 2006. In *Intergenerational transfers and social protection in Latin America*. Mexico City: National Transfers Accounts Working Paper Series.

Burchi, F. 2012. "Whose education affects a child's nutritional status? From parents' to household's education." *Demographic Research* 27:681-704. doi: 10.4054/DemRes.2012.27.23.

Cai, F., J. Giles, and X. Meng. 2006. "How well do children insure parents against low retirement income? An analysis using survey data from urban China." *Journal of Public Economics* 90 (12):2229-2255. doi: 10.1016/j.jpubeco.2006.03.004.

Caldwell, J. C. 1980. "Mass Education as a Determinant of the Timing of Fertility Decline." *Population and Development Review* 6 (2):225-255. doi: 10.2307/1972729.

Cameron, L. A., and D. Cobb-Clark. 2006. "Do coresidency and financial transfers from the children reduce the need for elderly parents to works in developing countries?" *Journal of Population Economics* 21 (4):1007-1033. doi: 10.1007/s00148-006-0105-8.

Case, A, and Ch Paxson. 2013. "HIV Risk and Adolescent Behaviors in Africa." *American Economic Review* 103 (3):433-438. doi: 10.1257/aer.103.3.433.

Case, A., D. Lubotsky, and C. Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review* 92 (5):1308-1334. doi: 10.1257/000282802762024520.

Central Statistics Office, Government of Botswana. 2001. Population and Housing Census. Private Bag 0024 Gaborone: Central Statistics Office Botswana.

Central Statistics Office, Government of Botswana. 2004. "2004 Botswana AIDS Impact Survey II."

Central Statistics Office, Government of Botswana. 2008a. 2005/06 Labour Force Report. edited by Department of Printing and Publishing Services. Gaborone: Central Statistics Office.

Central Statistics Office, Government of Botswana. 2008b. "2008 Botswana AIDS Impact Survey III."

Central Statistics Office, Government of Botswana. 2011. Population and Housing Census. Private Bag 0024 Gaborone: Central Statistics Office.

Chicoine, L. 2012. "Education and Fertility: Evidence from a Policy Change in Kenya." *Institute for the Study of Labor (IZA), Research Paper Series* 6778

Cho, H., D. D. Hallfors, I. I. Mbai, J. Itindi, B. W. Milimo, C. T. Halpern, and B. J. Iritani. 2011. "Keeping Adolescent Orphans in School to Prevent Human Immunodeficiency Virus Infection: Evidence From a Randomized Controlled Trial in Kenya." *Journal of Adolescent Health* 48 (5):523-526. doi: DOI 10.1016/j.jadohealth.2010.08.007.

Chou, K. L. 2008. "Parental Repayment Hypothesis in Intergenerational Financial Transfers from Adult Children to Elderly Parents: Evidence from Hong Kong." *Educational Gerontology* 34 (9):788-799. doi: 10.1080/03601270802095972.

Christiaensen, L., and H. Alderman. 2004. "Child Malnutrition in Ethiopia: Can Maternal Knowledge Augment the Role of Income?*." *Economic Development and Cultural Change* 52 (2):287-312. doi: 10.1086/380822.

Clark, D., and H. Royer. 2013. "The Effect of Education on Adult Mortality and Health: Evidence from Britain." *American Economic Review* 103 (6):2087-2120. doi: DOI 10.1257/aer.103.6.2087.

Cleland, J. G., and J. K. van Ginneken. 1988. "Maternal education and child survival in developing countries: The search for pathways of influence." *Social Science & Medicine* 27 (12):1357-1368. doi: 10.1016/0277-9536(88)90201-8.

Cochrane, S. H., J. Leslie, and D. J. O'Hara. 1982. "Parental education and child health: intracountry evidence." *Health Policy and Education* 2 (3-4):213-50.

Corsi, D. J., I. Mejía-Guevara, and S. V. Subramanian. 2015. "Risk factors for chronic undernutrition among children in India: Estimating relative importance, population attributable risk and fractions." *Social Science & Medicine* 157:165-85. doi: 10.1016/j.socscimed.2015.11.014.

Coyle, K. K., D. B. Kirby, B. V. Marin, C. A. Gomez, and S. E. Gregorich. 2004. "Draw the line/respect the line: a randomized trial of a middle school intervention to reduce sexual risk behaviors." *American Journal of Public Health* 94 (5):843-51.

Coyle, K. K., D. B. Kirby, L. E. Robin, S. W. Banspach, E. Baumler, and J. R. Glassman. 2006. "All4You! A randomized trial of an HIV, other STDs, and pregnancy prevention intervention for alternative school students." *AIDS Education and Prevention* 18 (3):187-203. doi: 10.1521/aeap.2006.18.3.187.

Crews, F., J. He, and C. Hodge. 2007. "Adolescent cortical development: A critical period of vulnerability for addiction." *Pharmacology Biochemistry and Behavior* 86 (2):189-199. doi: 10.1016/j.pbb.2006.12.001.

Cunha, F., and J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2):31-47. doi: 10.1257/aer.97.2.31.

Currie, J. 2009. "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development." *Journal of Economic Literature* 47 (1):87-122. doi: 10.1257/jel.47.1.87.

Currie, J., and E. Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *The Quarterly Journal of Economics* 118 (4):1495-1532. doi: 10.1162/003355303322552856.

Cutler, D. M., A. Deaton, and A. Lleras-Muney. 2006. "The Determinants of Mortality." *Journal of Economic Perspectives* 20 (3):97-120. doi: 10.1257/jep.20.3.97.

Cutler, D. M., and A. Lleras-Muney. 2006. "Education and Health: Evaluating Theories and Evidence." *National Bureau of Economic Research Working Paper Series* No. 12352. doi: 10.3386/w12352.

Cutler, D. M., and A. Lleras-Muney. 2010. "Understanding differences in health behaviors by education." *Journal of Health Economics* 29 (1):1-28. doi: 10.1016/j.jhealeco.2009.10.003.

Dahl, R. E. 2004. "Adolescent Brain Development: A Period of Vulnerabilities and Opportunities. Keynote Address." *Annals of the New York Academy of Sciences* 1021 (1):1-22. doi: 10.1196/annals.1308.001.

Dallabetta, G. A., P. G. Miotti, J. D. Chiphangwi, A. J. Saah, G. Liomba, N. Odaka, F. Sungani, and D. R. Hoover. 1993. "High Socioeconomic-Status Is a Risk Factor for Human-Immunodeficiency-Virus Type-1 (Hiv-1) Infection but Not for Sexually-Transmitted Diseases in Women in Malawi - Implications for Hiv-1 Control." *Journal of Infectious Diseases* 167 (1):36-42.

De Neve, J. W., G. Fink, S. V. Subramanian, S. Moyo, and J. Bor. 2015. "Length of secondary schooling and risk of HIV infection in Botswana: evidence from a natural experiment." *The Lancet Global Health* 3 (8):e470-7. doi: 10.1016/S2214-109X(15)00087-X.

De Neve, J. W., and S. V. Subramanian. 2017. "Causal Effect of Parental Schooling on Early Childhood Undernutrition — Quasi-experimental Evidence from Zimbabwe." *American Journal of Epidemiology* In press.

De Neve, J.W., and G. Harling. 2017. "Offspring schooling associated with increased parental survival in rural KwaZulu-Natal, South Africa." *Social Science & Medicine* 176:149-157. doi: 10.1016/j.socscimed.2017.01.015.

De Neve, J.W., and I. Kawachi. 2017. "Spillovers between siblings and from offspring to parents are understudied: A review and future directions for research." *Social Science & Medicine*. doi: 10.1016/j.socscimed.2017.04.010.

de Walque, D. 2009. "Does Education Affect HIV Status? Evidence from five African Countries." *World Bank Economic Review* 23 (2):209-233. doi: Doi 10.1093/Wber/Lhp005.

Deaton, A. 1997. *The analysis of household surveys : a microeconometric approach to development policy*. Baltimore, MD: Published for the World Bank by Johns Hopkins University Press.

Demographic and Health Surveys. 2011. DHS final reports Zimbabwe. Calverton.

DiClemente, R. J., G. M. Wingood, E. S. Rose, J. M. Sales, D. L. Lang, A. M. Caliendo, J. W. Hardin, and R. A. Crosby. 2009. "Efficacy of sexually transmitted disease/human immunodeficiency virus sexual risk-reduction intervention for african american adolescent females seeking sexual health services: a randomized controlled trial." *Archives of Pediatrics and Adolescent Medicine* 163 (12):1112-21. doi: 10.1001/archpediatrics.2009.205.

Dinçer, M. A., N. Kaushal, and M. Grossman. 2014. "Women's Education: Harbinger of Another Spring? Evidence from a Natural Experiment in Turkey." *World Development* 64:243-258. doi: 10.1016/j.worlddev.2014.06.010.

Duflo, E. 2001. "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment." *American Economic Review* 91 (4):795-813. doi: Doi 10.1257/Aer.91.4.795.

Duflo, E., P. Dupas, and M. Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review* 105 (9):2757-97. doi: 10.1257/aer.20121607.

Dunkle, K. L., R. K. Jewkes, H. C. Brown, G. E. Gray, J. A. McIntryre, and S. D. Harlow. 2004. "Transactional sex among women in Soweto, South Africa: prevalence, risk factors and association with HIV infection." *Social Science & Medicine* 59 (8):1581-92. doi: 10.1016/j.socscimed.2004.02.003.

Egata, G., Y. Berhane, and A. Worku. 2014. "Predictors of acute undernutrition among children aged 6 to 36 months in east rural Ethiopia: a community based nested case - control study." *BMC Pediatrics* 14 (1):91. doi: 10.1186/1471-2431-14-91.

Epstein, E., and R. Guttman. 1984. "Mate selection in man: evidence, theory, and outcome." *Social Biology* 31 (3-4):243-78.

Evans, R., and S. O. Becker. 2009. *Children caring for parents with HIV and AIDS : global issues and policy responses*. Bristol, UK: Policy Press.

Field, E, and J de la Roca. 2005. "Are there upward intergenerational education spillovers? The influence of children's schooling on parents' smoking cessation." Harvard University.

Finlay, J. E., E. Ozaltin, and D. Canning. 2011. "The association of maternal age with infant mortality, child anthropometric failure, diarrhoea and anaemia for first births: evidence from 55 low- and middle-income countries." *BMJ Open* 1 (2):e000226-e000226. doi: 10.1136/bmjopen-2011-000226.

Folbre, N. 1994. "Children as Public Goods." *The American Economic Review* 84 (2):86-90.

Fortson, J. G. 2008. "The gradient in sub-Saharan Africa: socioeconomic status and HIV/AIDS." *Demography* 45 (2):303-22.

Frankenberg, E., L. Lillard, and R. J. Willis. 2002. "Patterns of Intergenerational Transfers in Southeast Asia." *Journal of Marriage and Family* 64 (3):627-641. doi: 10.1111/j.1741-3737.2002.00627.x.

Friedman, E. M., and R. D. Mare. 2014. "The schooling of offspring and the survival of parents." *Demography* 51 (4):1271-93. doi: 10.1007/s13524-014-0303-z.

Frost, M. B., R. Forste, and D. W. Haas. 2005. "Maternal education and child nutritional status in Bolivia: finding the links." *Social Science & Medicine* 60 (2):395-407. doi: 10.1016/j.socscimed.2004.05.010.

Fylkesnes, K., R. M. Musonda, M. Sichone, Z. Ndhlovu, F. Tembo, and M. Monze. 2001. "Declining HIV prevalence and risk behaviours in Zambia: evidence from surveillance and population-based surveys." *AIDS* 15 (7):907-16.

Gakidou, E., K. Cowling, R. Lozano, and C. J. L. Murray. 2010. "Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis." *The Lancet* 376 (9745):959-974. doi: 10.1016/s0140-6736(10)61257-3.

Galea, S., M. Tracy, K. J. Hoggatt, C. DiMaggio, and A. Karpati. 2011. "Estimated Deaths Attributable to Social Factors in the United States." *American Journal of Public Health* 101 (8):1456-1465. doi: 10.2105/ajph.2010.300086.

Gaotlhobogwe, M. 2010. "Attitudes to and perceptions of design and technology students towards the subject: a case of five junior secondary schools in Botswana." *Thesis, University of Wales*:354.

Giles, J., D. Wang, and C. Zhao. 2011. "Can China's Rural Elderly Count on Support from Adult Children? Implications of Rural-to-Urban Migration." *Journal of Population Ageing* 3 (3-4):183-204. doi: 10.1007/s12062-011-9036-6.

Glewwe, P. 1999. "Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco." *The Journal of Human Resources* 34 (1):124-159. doi: 10.2307/146305.

Glied, S., and A. Lleras-Muney. 2008. "Technological innovation and inequality in health." *Demography* 45 (3):741-61.

Government of Botswana. 1994. "The revised national policy on education." Botswana Government Printer.

Government of Botswana. 1995. "Curriculum Blueprint: Ten year Basic Education Programme."

Government of Botswana. 2009. *Botswana AIDS impact survey III : statistical report*. Gaborone: Central Statistics Office.

Government of Botswana. 2013. "Education Report Drawn From 2009/10 Botswana Core Welfare Indicators Survey." Central Statistics Office Botswana Accessed 2013/15.

Grépin, K. A., and P. Bharadwaj. 2015. "Maternal education and child mortality in Zimbabwe." *Journal of Health Economics* 44:97-117. doi: 10.1016/j.jhealeco.2015.08.003.

Grossman, M. 2006. "Chapter 10 Education and Nonmarket Outcomes." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch, 577-633. Amsterdam: North Holland.

Grossman, M. 2015. "The Relationship between Health and Schooling: What's New?" *NBER Working Paper No. 21609*. doi: 10.3386/w21609.

Group, WHO Multicentre Growth Reference Study. 2006. WHO child growth standards: length/ height-for-age, weight for-age, weight-for-length and body mass index for age: methods and development. edited by WHO Multicentre Growth Reference Study Group. Geneva: World Health Organization.

Güneş, P. M. 2015. "The role of maternal education in child health: Evidence from a compulsory schooling law." *Economics of Education Review* 47:1-16. doi: 10.1016/j.econedurev.2015.02.008.

Hargreaves, J. R., C. P. Bonell, T. Boler, D. Boccia, I. Birdthistle, A. Fletcher, P. M. Pronyk, and J. R. Glynn. 2008. "Systematic review exploring time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa." *AIDS* 22 (3):403-14. doi: 10.1097/QAD.0b013e3282f2aac3.

Harling, G., M. L. Newell, F. Tanser, I. Kawachi, S. V. Subramanian, and T. Bärnighausen. 2014. "Do Age-Disparate Relationships Drive HIV Incidence in Young Women? Evidence from a Population Cohort in Rural KwaZulu-Natal, South Africa." *JAIDS Journal of Acquired Immune Deficiency Syndromes* 66 (4):443-451. doi: 10.1097/qai.0000000000000198.

Hernan, M. A., and J. M. Robins. 2006. "Instruments for causal inference: an epidemiologist's dream?" *Epidemiology* 17 (4):360-72. doi: 10.1097/01.ede.0000222409.00878.37.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2):467. doi: 10.2307/2951620.

Iorio, D., and R. Santaeulalia-Llopis. 2011. "Education, HIV Status, and Risky Sexual Behavior: How Much Does the Stage of the HIV Epidemic Matter?" *Working Paper, Universitat Autonoma de Barcelona*.

IPUMS. 2014. Integrated Public Use Microdata Series, International: Version 6.3 [Machine-readable database]. edited by University of Minnesota. Minneapolis.

Jacques Be-Ofuriyua, E. 2011. "Maternal education and child nutritional status in the Democratic Republic of Congo." *Journal of Public Health and Epidemiology* 3 (12). doi: 10.5897/jphe11.130.

Jemmott, J. B., 3rd, L. S. Jemmott, P. K. Braverman, and G. T. Fong. 2005. "HIV/STD risk reduction interventions for African American and Latino adolescent girls at an adolescent medicine clinic: a randomized controlled trial." *Archives of Pediatrics and Adolescent Medicine* 159 (5):440-9. doi: 10.1001/archpedi.159.5.440.

Jemmott, J. B., 3rd, L. S. Jemmott, and G. T. Fong. 2010. "Efficacy of a theory-based abstinence-only intervention over 24 months: a randomized controlled trial with young adolescents." *Archives of Pediatrics and Adolescent Medicine* 164 (2):152-9. doi: 10.1001/archpediatrics.2009.267.

Joint United Nations Programme on HIV/AIDS. 2013. Global report: UNAIDS report on the global AIDS epidemic 2013. Geneva, Switzerland.

Jukes, M., S. Simmons, and D. Bundy. 2008. "Education and vulnerability: the role of schools in protecting young women and girls from HIV in southern Africa." *AIDS* 22 Suppl 4:S41-56. doi: 10.1097/01.aids.0000341776.71253.04.

Kaufman, J. S., and R. S. Cooper. 1999. "Seeking causal explanations in social epidemiology." *American Journal of Epidemiology* 150 (2):113-20.

Kirby, D. B. 2008. "The impact of abstinence and comprehensive sex and STD/HIV education programs on adolescent sexual behavior." *Sexuality Research and Social Policy* 5 (3):18-27. doi: 10.1525/srsp.2008.5.3.18.

Kirunga, C. T., and J. P. Ntozi. 1997. "Socio-economic determinants of HIV serostatus: a study of Rakai District, Uganda." *Health Transition Review* 7 Suppl:175-88.

Knodel, J., J. Friedman, T. S. Anh, and B. T. Cuong. 2000. "Intergenerational Exchanges in Vietnam: Family Size, Sex Composition, and the Location of Children." *Population Studies* 54 (1):89-104.

Kohler, H. P., and R. L. Thornton. 2011. "Conditional Cash Transfers and HIV/AIDS Prevention: Unconditionally Promising?" *The World Bank Economic Review* 26 (2):165-190. doi: 10.1093/wber/lhr041.

Kohler, I. V., H. P. Kohler, P. Anglewicz, and J. R. Behrman. 2012. "Intergenerational transfers in the era of HIV/AIDS: Evidence from rural Malawi." *Demographic Research* 27:775-834.

Koniak-Griffin, D., J. Lesser, A. Nyamathi, G. Uman, J. A. Stein, and W. G. Cumberland. 2003. "Project CHARM: an HIV prevention program for adolescent mothers." *Family & Community Health* 26 (2):94-107.

Krapohl, E., K. Rimfeld, N. G. Shakeshaft, M. Trzaskowski, A. McMillan, J. B. Pingault, K. Asbury, N. Harlaar, Y. Kovas, P. S. Dale, and R. Plomin. 2014. "The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence." *Proceedings of the National Academy of Sciences* 111 (42):15273-15278. doi: 10.1073/pnas.1408777111.

Kravdal, O. 2002. "Education and Fertility in Sub-Saharan Africa." *Demography* 39 (2):233-250. doi: 10.1353/dem.2002.0017.

Krieger, N. 2001. "Theories for social epidemiology in the 21st century: an ecosocial perspective." *International Journal of Epidemiology* 30 (4):668-677. doi: 10.1093/ije/30.4.668.

Kuh, D. 2003. "Life course epidemiology." *Journal of Epidemiology & Community Health* 57 (10):778-783. doi: 10.1136/jech.57.10.778.

Kuziemko, I. 2014. "Human Capital Spillovers in Families: Do Parents Learn from or Lean on Their Children?" *Journal of Labor Economics* 32 (4):755-786. doi: 10.1086/677231.

Kuziemko, I., and J. Ferrie. 2014. "The Role of Immigrant Children in Their Parents' Assimilation in the United States, 1850-2010." In *Human Capital in History*.

Leclerc-Madlala, S. 2008. "Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability." *AIDS* 22 Suppl 4:S17-25. doi: 10.1097/01.aids.0000341774.86500.53.

Lee, R. D., and A. Mason. 2011. *Population aging and the generational economy : a global perspective*. Cheltenham ; Northampton, MA.

Levinsohn, J., T. Dinkelman, and R. Majelantle. 2006. "When Knowledge is not Enough: HIV/AIDS Information and Risky Behavior in Botswana." *National Bureau of Economic Research Working Paper Series* No. 12418. doi: 10.3386/w12418.

Lillard, L. A., and R. J. Willis. 1997. "Motives for Intergenerational Transfers: Evidence from Malaysia." *Demography* 34 (1):115. doi: 10.2307/2061663.

Lim, S. S., T. Vos, A. D. Flaxman, G. Danaei, K. Shibuya, H. Adair-Rohani, M. A. AlMazroa, M. Amann, H. R. Anderson, K. G. Andrews, M. Aryee, C. Atkinson, L. J. Bacchus, A. N. Bahalim, K. Balakrishnan, J. Balmes, S. Barker-Collo, A. Baxter, M. L. Bell, J. D. Blore, F. Blyth, C. Bonner, G. Borges, R. Bourne, M. Boussinesq, M.l Brauer, P. Brooks, N. G. Bruce, B. Brunekreef, C. Bryan-Hancock, C. Bucello, R. Buchbinder, F. Bull, R. T. Burnett, T. E. Byers, B. Calabria, J. Carapetis, E. Carnahan, Z. Chafe, F. Charlson, H. Chen, J. S. Chen, A. T. A. Cheng, J. C. Child, A. Cohen, K. E. Colson, B. C. Cowie, S. Darby, S. Darling, A. Davis, L. Degenhardt, F. Dentener, D. C. Des Jarlais, K. Devries, M. Dherani, E. L. Ding, E. R. Dorsey, T. Driscoll, K. Edmond, S. E. Ali, R. E. Engell, P. J. Erwin, S. Fahimi, G. Falder, F. Farzadfar, A. Ferrari, M. M. Finucane, S. Flaxman, F. G. R. Fowkes, G. Freedman, M. K. Freeman, E. Gakidou, S. Ghosh, E. Giovannucci, G. Gmel, K. Graham, R. Grainger, B. Grant, D. Gunnell, H. R. Gutierrez, W. Hall, H. W. Hoek, A. Hogan, H. D. Hosgood, D. Hoy, H. Hu, B. J. Hubbell, S. J. Hutchings, S. E. Ibeanusi, G. L. Jacklyn, R. Jasrasaria, J. B. Jonas, H. Kan, J. A. Kanis, N. Kassebaum, N. Kawakami, Y. H. Khang, S. Khatibzadeh, J. P. Khoo, C. Kok, F. Laden, R. Lalloo, Q. Lan, T. Lathlean, J. L. Leasher, J. Leigh, Y. Li, J. K. Lin, S. E. Lipshultz, S. London, R. Lozano, Y. Lu, J. Mak, R. Malekzadeh, L. Mallinger, W. Marcenes, L. March, R. Marks, R. Martin, P. McGale, J. McGrath, S.i Mehta, Z. A. Memish, G. A. Mensah, T. R. Merriman, R. Micha, C. Michaud, V. Mishra, K. M. Hanafiah, A. A. Mokdad, L. Morawska, D. Mozaffarian, T. Murphy, M. Naghavi, B. Neal, P. K. Nelson, J. M. Nolla, R. Norman, C. Olives, S. B. Omer, J. Orchard, R. Osborne, B. Ostro, A. Page, K. D. Pandey, C. D. H. Parry, E. Passmore, J. Patra, N. Pearce, P. M. Pelizzari, M. Petzold, M. R. Phillips, D. Pope, C. A. Pope, J. Powles, M. Rao, H. Razavi, E. A. Rehfuess, J. T. Rehm, B. Ritz, F. P. Rivara, T. Roberts, C. Robinson, J. A. Rodriguez-Portales, I. Romieu, R. Room, L. C. Rosenfeld, A. Roy, L. Rushton, J. A. Salomon, U. Sampson, L. Sanchez-Riera, E. Sanman, A. Sapkota, S. Seedat, P. Shi, K. Shield, R. Shivakoti, G. M. Singh, D. A. Sleet, E. Smith, K. R. Smith, N. J. C. Stapelberg, K. Steenland, H. Stöckl, L. J. Stovner, K. Straif, L. Straney, G. D. Thurston, J. H. Tran, R. Van Dingenen, A. van Donkelaar, J. L. Veerman, L. Vijayakumar, R. Weintraub, M. M. Weissman, R. A. White, H. Whiteford, S. T. Wiersma, J. D. Wilkinson, H. C. Williams, W. Williams, N. Wilson, A. D. Woolf, P. Yip, J. M. Zielinski, A. D. Lopez, C. J. L. Murray, and M. Ezzati. 2012. "A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010." *The Lancet* 380 (9859):2224-2260. doi: 10.1016/s0140-6736(12)61766-8.

Lindeboom, M., A. Llena-Nozal, and B. van der Klaauw. 2009. "Parental education and child health: Evidence from a schooling reform." *Journal of Health Economics* 28 (1):109-131. doi: 10.1016/j.jhealeco.2008.08.003.

Little, R. J., and D. B. Rubin. 2000. "Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches." *Annual Review of Public Health* 21:121-45. doi: 10.1146/annurev.publhealth.21.1.121.

Lleras-Muney, A. 2005. "The Relationship Between Education and Adult Mortality in the United States." *Review of Economic Studies* 72 (1):189-221. doi: 10.1111/0034-6527.00329.

Longfield, K., A. Glick, M. Waithaka, and J. Berman. 2004. "Relationships Between Older Men and Younger Women: Implications for STIs/HIV in Kenya." *Studies in Family Planning* 35 (2):125-134. doi: 10.1111/j.1728-4465.2004.00014.x.

Luke, N. 2005. "Confronting the 'sugar daddy' stereotype: age and economic asymmetries and risky sexual behavior in urban Kenya." *International Family Planning Perspectives* 31 (1):6-14. doi: 10.1363/ifpp.31.06.05.

Lundborg, P., and K. Majlesi. 2015. "Intergenerational Transmission of Human Capital: Is It a One-Way Street?" *Institute for the Study of Labor Discussion Paper No. 9280*.

Mabsout, R., and I. van Staveren. 2010. "Disentangling Bargaining Power from Individual and Household Level to Institutions: Evidence on Women's Position in Ethiopia." *World Development* 38 (5):783-796. doi: 10.1016/j.worlddev.2009.11.011.

Makoka, D., and P. K. Masibo. 2015. "Is there a threshold level of maternal education sufficient to reduce child undernutrition? Evidence from Malawi, Tanzania and Zimbabwe." *BMC Pediatrics* 15 (1). doi: 10.1186/s12887-015-0406-8.

Marmot, M. G. 2003. "Understanding social inequalities in health." *Perspectives in Biology and Medicine* 46 (3 Suppl):S9-23.

Martin, T. C. 1995. "Women's Education and Fertility: Results from 26 Demographic and Health Surveys." *Studies in Family Planning* 26 (4):187-202. doi: 10.2307/2137845.

Mason, A., and R. D. Lee. 2006. "Reform and support systems for the elderly in developing countries: capturing the second demographic dividend." *Genus* 62 (2):11-35. doi: 10.2307/29789308.

McCrary, J., and H. Royer. 2011. "The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth." *American Economic Review* 101 (1):158-195. doi: DOI 10.1257/aer.101.1.158.

McGarry, K. 1998. "Caring for the elderly: The role of adult children." *Inquiries in the Economics of Aging*:133-166.

Meara, E. R., S. Richards, and D. M. Cutler. 2008. "The Gap Gets Bigger: Changes In Mortality And Life Expectancy, By Education, 1981-2000." *Health Affairs* 27 (2):350-360. doi: 10.1377/hlthaff.27.2.350.

Mincer, J. 1974. Schooling, Experience, and Earnings. New York: NBER.

Monstad, Karin, Carol Propper, and Kjell G. Salvanes. 2008. "Education and Fertility: Evidence from a Natural Experiment*." *Scandinavian Journal of Economics* 110 (4):827-852. doi: 10.1111/j.1467-9442.2008.00563.x.

Nyerere, J.K. 1967. Education for Self Reliance. In *Education for Self Reliance*. United Republic of Tanzania, Arusha.

Oketch, M., and C. Rolleston. 2007. "Chapter 5 Policies on Free Primary and Secondary Education in East Africa: Retrospect and Prospect." *Review of Research in Education* 31 (1):131-158. doi: 10.3102/0091732x07300046131.

Omari, I. M. 1983. *Universal Primary Education in Tanzania*. Edited by Ottawa: International Development Research Centre (IDRC-TS42e).

Osili, U. O., and Br. T. Long. 2008. "Does female schooling reduce fertility? Evidence from Nigeria." *Journal of Development Economics* 87 (1):57-75. doi: 10.1016/j.jdeveco.2007.10.003.

Parliament of Tanzania. 1977. *The National Education Act*.

Penny, M. E., H. M. Creed-Kanashiro, R. C. Robert, M. R. Narro, L. E. Caulfield, and R. E. Black. 2005. "Effectiveness of an educational intervention delivered through the health services to improve nutrition in young children: a cluster-randomised controlled trial." *The Lancet* 365 (9474):1863-1872. doi: 10.1016/s0140-6736(05)66426-4.

Pettifor, A. E., H. V. Rees, I. Kleinschmidt, A. E. Steffenson, C. MacPhail, L. Hlongwa-Madikizela, K. Vermaak, and N. S. Padian. 2005. "Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey." *AIDS* 19 (14):1525-34.

Pettifor, A. E., A. van der Straten, M. S. Dunbar, S. C. Shiboski, and N. S. Padian. 2004. "Early age of first sex: a risk factor for HIV infection among women in Zimbabwe." *AIDS* 18 (10):1435-1442. doi: DOI 10.1097/01.aids.0000131338.61042.b8.

Pettifor, A; , C; MacPhail, A; Selin, X; Gomez-Olivé, J; Hughes, R; Wagner, and W; Mokoena Mabuza, I; Eshleman, S; Piwowar-Manning, E; Twine, R; Julien, A; Marcus, C; Andrew, P; Wang, J; Xing, Y; McKinstry, L; Hamilton, E; Agyei, Y; Allison, S; Sato, P; Townley, E; Tollman, S; Kahn, K; HPTN 068 Study Team. 2015. "HPTN 068 conditional cash transfer to prevent HIV infection among young women in South Africa: results of a randomized controlled trial." *Journal of the International AIDS Society* 18 (5Suppl 4).

Phiri, T. 2014. "Review of Maternal Effects on Early Childhood Stunting." *Grand Challenges Canada Working Paper Series, GCC 14-09*.

Pollak, R. 2005. "Bargaining Power in Marriage: Earnings, Wage Rates and Household Production." *National Bureau of Economic Research Working Paper Series* No. 11239. doi: 10.3386/w11239.

Psacharopoulos, G. 1994. "Returns to Investment in Education - a Global Update." *World Development* 22 (9):1325-1343. doi: Doi 10.1016/0305-750x(94)90007-8.

Rende, S. 2014. "The Relationship Between a Mother's Attitude Toward Domestic Violence and Children's Schooling Outcomes in Turkey." *Journal of Interpersonal Violence* 29 (14):2548-2570. doi: 10.1177/0886260513520468.

Robert, E. B. L., and O. Stark. 1985. "Motivations to Remit: Evidence from Botswana." *Journal of Political Economy* 93 (5):901-918.

Rosenberg, M., A. Pettifor, W. C. Miller, H. Thirumurthy, M. Emch, S. A. Afolabi, K. Kahn, M. Collinson, and S. Tollman. 2015. "Relationship between school dropout and teen pregnancy among rural South African young women." *International Journal of Epidemiology* 44 (3):928-36. doi: 10.1093/ije/dyv007.

Rosenzweig, M. R., and T. P. Schultz. 1982. "Child mortality and fertility in Colombia: Individual and community effects." *Health Policy and Education* 2 (3-4):305-348. doi: 10.1016/0165-2281(82)90015-7.

Ross, C. E., and C. L. Wu. 1995. "The Links Between Education and Health." *American Sociological Review* 60 (5):719-745. doi: 10.2307/2096319.

Rowa-Dewar, N., A. Amos, and S. Cunningham-Burley. 2014. "Children's resistance to parents' smoking in the home and car: a qualitative study." *Addiction* 109 (4):645-52. doi: 10.1111/add.12435.

Ruel, M. T., and H. Alderman. 2013. "Nutrition-sensitive interventions and programmes: how can they help to accelerate progress in improving maternal and child nutrition?" *The Lancet* 382 (9891):536-551. doi: 10.1016/s0140-6736(13)60843-0.

Saad, P. M. 2005. "Informal support transfers of the elderly in Brazil and Latin America." In *Sixty plus: The elderly Brazilians and their new social roles*. Rio de Janeiro: Institute of Applied Economic Research.

Sabater, Albert, and Elspeth Graham. 2016. "The Role of Children's Education for the Mental Health of Aging Migrants in Europe." *GeroPsych* 29 (2):81-92. doi: 10.1024/1662-9647/a000145.

Sabates, R, J Westbrook, and J Hernandez-Fernandez. 2011. The health and education benefits of universal primary education for the next generation: Evidence from Tanzania. In *CREATE Pathways to Access: Research Monograph No.62*, edited by University of Sussex.

Sabates, R., J. Westbrook, and J. Hernandez-Fernandez. 2012. "The 1977 Universal Primary Education in Tanzania: a historical base for quantitative enquiry." *International Journal of Research & Method in Education* 35 (1):55-70. doi: 10.1080/1743727x.2011.609551.

Sackey, H. A. 2005. *Female labour force participation in Ghana : the effects of education*, *AERC research paper*. Nairobi: African Economic Research Consortium.

Schultz, T. W. 1960. "Capital Formation by Education." *Journal of Political Economy* 68 (6):571-583. doi: Doi 10.1086/258393.

Sen, A. 1980. Sen A. Equality of What?. In: McMurrin S Tanner Lectures on Human Values, Volume 1. Cambridge: Cambridge University Press ; 1980. In *McMurrin S Tanner Lectures on Human Values*: Cambridge University Press.

Shrier, L. A., R. Ancheta, E. Goodman, V. M. Chiou, M. R. Lyden, and S. J. Emans. 2001. "Randomized controlled trial of a safer sex intervention for high-risk adolescent girls." *Archives of Pediatrics and Adolescent Medicine* 155 (1):73-9.

Sikkema, K. J., E. S. Anderson, J. A. Kelly, R. A. Winett, C. Gore-Felton, R. A. Roffman, T. G. Heckman, K. Graves, R. G. Hoffmann, and M. J. Brondino. 2005. "Outcomes of a randomized, controlled community-level HIV prevention intervention for adolescents in low-income housing developments." *AIDS* 19 (14):1509-16.

Silvestrin, S., C. H. da Silva, V. N. Hirakata, A. A. S. Goldani, P. P. Silveira, and M. Z. Goldani. 2013. "Maternal education level and low birth weight: a meta-analysis." *Jornal de Pediatria* 89 (4):339-345. doi: 10.1016/j.jped.2013.01.003.

Siphambe, H. K. 2000. "Rates of return to education in Botswana." *Economics of Education Review* 19 (3):291-300. doi: Doi 10.1016/S0272-7757(99)00042-4.

Steinberg, L. 2005. "Cognitive and affective development in adolescence." *Trends in Cognitive Sciences* 9 (2):69-74. doi: 10.1016/j.tics.2004.12.005.

The World Bank. 2015. "World Development Indicators."

Torssander, J. 2013. "From child to parent? The significance of children's education for their parents' longevity." *Demography* 50 (2):637-59. doi: 10.1007/s13524-012-0155-3.

Tortolero, S. R., C. M. Markham, M. F. Peskin, R. Shegog, R. C. Addy, S. L. Escobar-Chaves, and E. R. Baumler. 2010. "It's Your Game: Keep It Real: delaying sexual behavior with an effective middle school program." *Journal of Adolescent Health* 46 (2):169-79. doi: 10.1016/j.jadohealth.2009.06.008.

Tsheko, G. 2012. "HIV/AIDS and Education in Botswana." In *The Impact of HIV/AIDS on Education Worldwide (International Perspectives on Education and Society, Volume 18)*, edited by Ryan N. Glover Alexander W. Wiseman, 141-159. Emerald Group Publishing Limited.

UNESCO Institute for Statistics. 2015. "Tanzania Education Statistics."

Vandemoortele, J., and E. Delamonica. 2000. "The 'education vaccine' against HIV." *Current Issues in Comparative Education* 3 (1).

Vazir, S., P. Engle, N. Balakrishna, P. L. Griffiths, S. L. Johnson, H. M. Creed-Kanashiro, S. Fernandez Rao, M. R. Shroff, and M. E. Bentley. 2013. "Cluster-randomized trial on complementary and responsive feeding education to caregivers found improved dietary intake, growth and development among rural Indian toddlers." *Maternal & Child Nutrition* 9 (1):99-117. doi: 10.1111/j.1740-8709.2012.00413.x.

Victoria, C. G., S. R. A. Huttly, F. C. Barros, C. Lombardi, and J. P. Vaughan. 1992. "Maternal education in relation to early and late child health outcomes: Findings from a Brazilian cohort study." *Social Science & Medicine* 34 (8):899-905. doi: 10.1016/0277-9536(92)90258-r.

Vikram, K., R. Vanneman, and S. Desai. 2012. "Linkages between maternal education and childhood immunization in India." *Social Science & Medicine* 75 (2):331-339. doi: 10.1016/j.socscimed.2012.02.043.

Villarruel, A. M., J. B. Jemmott, 3rd, and L. S. Jemmott. 2006. "A randomized controlled trial testing an HIV prevention intervention for Latino youth." *Archives of Pediatrics and Adolescent Medicine* 160 (8):772-7. doi: 10.1001/archpedi.160.8.772.

Vollmer S, Bommer C, Krishna A, Harttgen K, and S. V. Subramanian. 2016. "The association of parental education with childhood undernutrition in low- and middle-income countries: comparing the role of paternal and maternal education." *International Journal of Epidemiology* In press. doi: 10.1093/ije/dyw133.

Wilde, E. T., L. Batchelder, and D. Ellwood. 2010. The Mommy Track Divides: The Impact of Childbearing on Wages of Women of Differing Skill Levels. In *NBER Working Paper No. 16582*.

Willis, R.J. 1973. "A new approach to the economic theory of fertility behavior." *Journal of Political Economy* 81 (2).

Wojcicki, J. M. 2014. "The double burden household in sub-Saharan Africa: maternal overweight and obesity and childhood undernutrition from the year 2000: results from World Health Organization Data (WHO) and Demographic Health Surveys (DHS)." *BMC Public Health* 14 (1):1124. doi: 10.1186/1471-2458-14-1124.

Wooldridge, J.M. 2013. *Introductory Econometrics: A Modern Approach*: Michigan State University.

Yahirun, J. J., C. M. Sheehan, and M. D. Hayward. 2016. "Adult Childrens Education and Parents Functional Limitations in Mexico." *Research on Aging* 38 (3):322-345. doi: 10.1177/0164027515620240.

Zimmer, Z., A. I. Hermalin, and H. S. Lin. 2002. "Whose education counts? The added impact of adult-child education on physical functioning of older taiwanese." *J Gerontol B Psychol Sci Soc Sci* 57 (1):S23-32.

Appendix A

Appendix to Chapter 1

A.1 Additional Robustness Checks

Using an Earlier Education Policy Reform

Previous school reforms led to changes in the distribution of schooling for cohorts born before 1975, so they were excluded from my main analysis. As a robustness check, I used a previous policy reform as an additional source of exogenous variation. In January 1986, Botswana temporarily shifted the tenth year of education from junior secondary to senior secondary school, where access was lower. This 1986 reform was identical in nature as the policy reform I exploit in my main analysis but occurred in the opposite direction (tenth grade was shifted *to* senior secondary school). As a result of both of these policy reforms, a decade-long '7+2+3' system existed consisting of only two years of junior secondary but three years of senior secondary school. Similar to my main analysis, I defined exposure to the absence of this reform period based on birth cohort and legal school entry age, and used this variable as an instrument for educational attainment (zero if born between 1971 and 1980, and one otherwise). Figure A2 shows trends across birth cohorts in the probability of completing either at least nine or ten years of schooling. The fraction of students completing at least nine years of schooling rose continuously across birth cohorts. However, the share of students with at least ten years of schooling decreased sharply for the birth cohorts affected by the reform period.

Falsification Test

I used South African survey data to assess HIV prevalence among cohorts of ethnic Tswana born and residing just across the border in South Africa, who were not exposed to the education policy reform of 1996. The Botswana – South Africa border is a colonial boundary created in 1885, when the southern portion of Bechuanaland was subsumed into British Cape Colony, while the northern portion formed the British Bechuanaland Protectorate. The majority of ethnic Tswana reside in North West Province in South Africa. I used HIV survey data on ethnic Tswana (Setswana-speakers), born in South Africa, ages at least 18 years old, born after 1975, and with a valid HIV test result in the 2002 and 2005 national HIV surveys. Similar to my main analysis for Botswana, I used the 1996 policy reform as an instrument for years of schooling. Figure A3 shows HIV prevalence by birth cohort for Batswana and South African ethnic Tswana separately. Among cohorts exposed to the current '7+3+2' regime, HIV infection rates by birth cohort are similar for Batswana and South African Tswana. However, there is a marked increase in HIV infection rates for Batswana cohorts exposed to the reform. To assess whether my models were merely picking up cohort effects associated with the timing of the epidemic in the region, I estimated my main regression model (equations 1, 2, in the main text) for ethnic Tswana born in and residing in South Africa, who were not exposed to the policy reform in Botswana.

Attrition and HIV Test Consent Rates

Results could be biased by differential HIV test consent rates by birth cohort or by selection bias from mortality risk associated with being born after 1980. Increased education may have improved access to ART, which became available in 2002 in Botswana; however, this would lead to a higher HIV prevalence among those with more education, and hence a bias against the direction of my results. I found no evidence of a post-1980 cohort effect in either of these variables that might bias my estimates. I also assessed whether differential non-consent might have biased my results, by imputing HIV status for respondents with missing HIV biomarker data and found similar results. Further, in a regression framework, consent rates did not change much with exposure to the reform (De Neve et al. 2015). Lastly, the datasets I used do not contain information on interviewer identity or time of interview, which would have allowed me to use Heckman-type selection models to correct for selection on unobserved variables (Bärnighausen et al. 2011).

Exclusion of 1980 Birth Cohort

In analysis of natural experiments, it is best practice to use the legal structure of the policy change as an instrument for the actual changes in schooling observed, since legal changes are exogenous to individuals' choices. The first cohort that would have been affected *de jure* was the 1981 birth cohort. However, I also observed increased levels of grade 10 completion in the 1980 birth cohort, which may have included individuals who misreported their age or years of schooling, or entered school late or repeated grades. I assessed the robustness of my results excluding respondents born in 1980, which were partially affected by the reform.

Weighting

The use of sample weights is subject to controversy in analytical inference. As an additional robustness check, I added sample weights to my main model.

Alternative Hypotheses Possibly Affecting Post-1980 Birth Cohorts.

I evaluated other social reforms or population changes that might have happened around the time of the education policy reform of 1996—and potentially have affected birth cohorts born after 1980—by conducting a desk review of both published and unpublished documents. These included the Revised National Policy of Education (1994), the Revised National Population Policy (1997), the Seventh National Development Plan (1991-1997), as well as articles on Botswana's education system and curriculum. I listed alternative policy hypotheses with a brief description of each policy and, based on the description, whether or not these could be a threat to my main identification strategy. Potential hypotheses included the introduction of HIV education in the formal school curriculum, the roll-out of ART, HIV prevention programs, abortion or family planning policy changes, socio-economic reforms, vaccine or disaster relief programs if these would have affected specific birth cohorts.

A.2 Results of Robustness Checks

Table A5 shows the robustness of the ITT and 2SLS results to additional controls by sex and in the pooled sample, including a quadratic and/or cubic term in age, quadratic term in year of birth, year of birth*surveywave interactions, sample weights, year of birth*reformindicator interactions, narrower birth cohort windows, imputed HIV estimates and a 2SLS difference-in-difference estimator. In the pooled sample, using a quadratic term in age, cubic spline in age, quadratic term in year of birth, or a quadratic in yearofbirth*surveywave interactions, those who stayed in secondary school for an additional year had a 7 percentage point lower risk of HIV infection ($p = 0.048$, $p = 0.041$, $p = 0.014$, and $p = 0.010$, respectively); 8 percentage points ($p = 0.008$ and $p = 0.025$, respectively) using yearofbirth*surveywave interactions or weights; 9 - 11 percentage points ($p = 0.045$ and $p = 0.036$, respectively) using imputed HIV estimates; 9 percentage points ($p = 0.004$) using a 2SLS difference-in-difference estimator. In the placebo test, I found that the effect of the policy change was observed only among those with at least 9 years of schooling. I found no effect of the policy change on people with less than nine years of schooling, who experienced no increase in schooling. Excluding the 1980 birth cohort, schooling remained protective against HIV infection risk (-0.052 percentage points, p = 0.035, pooled sample). Results in robustness checks by sex were otherwise similar to my main results above.

Table A6 presents regression results – first stage, reduced form, 2SLS – using a previous education policy reform which took place in 1986. The absence of the reform period induced an increase of 0.4 years of schooling in the pooled sample. Birth cohorts that were not exposed to the reform period were 8 percentage points (p < 0.001) less likely to be HIV positive. In 2SLS estimates, one additional year of secondary schooling induced a decrease in 20 percentage points (p = 0.010) in HIV risk in the pooled sample; 42 percentage points (p = 0.088) for women and 5 percentage points (p = 0.249) for men.

Table A7 presents regression results – first stage, OLS, reduced form, 2SLS – for South African Tswana as a falsification test. In the "naïve" OLS model, years of schooling was associated with a decrease in HIV infection probability, similar to Botswana. However, the instrument had no significant effect on schooling, or on HIV infection rates.

I found no evidence of reforms affecting the same birth cohorts as those affected by the 1996 reform (Table A8). For instance, the NCE also recommended a diversification of Botswana's school curriculum. The NCE's long term goal was to increase the vocational orientation of academic subjects, increase the number of practical subjects, emphasize professional skills, relate the curriculum to the professional environment, and increase career guidance. More emphasis was placed on developing practical and professional skills and formal links with the private sector (e.g., through internships). However, these recommendations by the NCE were scheduled to be implemented gradually and, in some cases, left to the discretion of schools themselves.

A.3 References for Appendix A

Bärnighausen, T., Bor, J., Wandira-Kazibwe, S., Canning, D. (2011). Correcting HIV prevalence estimates for survey nonparticipation using Heckman-type selection models. *Epidemiology*, 22(1):27-35.

Borkum, E. (2009). Grade structure, educational attainment and labor market outcomes: Evidence from Botswana. *Job Market Paper Columbia University*

Government of Botswana. Penal Code (Amendment) Act of 11 October 1991 (1991). Government Gazette, Supplement A, A55-6.

Government of Botswana (1995). Curriculum Blueprint: Ten year Basic Education Programme. Gaborone.

Government of Botswana (1994). The revised national policy on education. Gaborone, Botswana: Botswana Government Printer.

Government of Botswana (2007). Central Statistics Office. Labour force survey 2005/06. Gaborone: Government Printer.

Deaton, A. (1997). The analysis of household surveys: a microeconometric approach to development policy. Baltimore, MD: Published for the World Bank by Johns Hopkins University Press.

De Neve, J.W., Fink, G., Subramanian, S.V., Moyo, S., Bor, J. (2015). Length of secondary schooling and risk of HIV infection in Botswana: evidence from a natural experiment. *The Lancet Global Health*. 3(8):e470-7.

Durrleman, S., Simon, R. (1989). Flexible regression models with cubic splines. *Statistics in Medicine*, 8(5):551-61

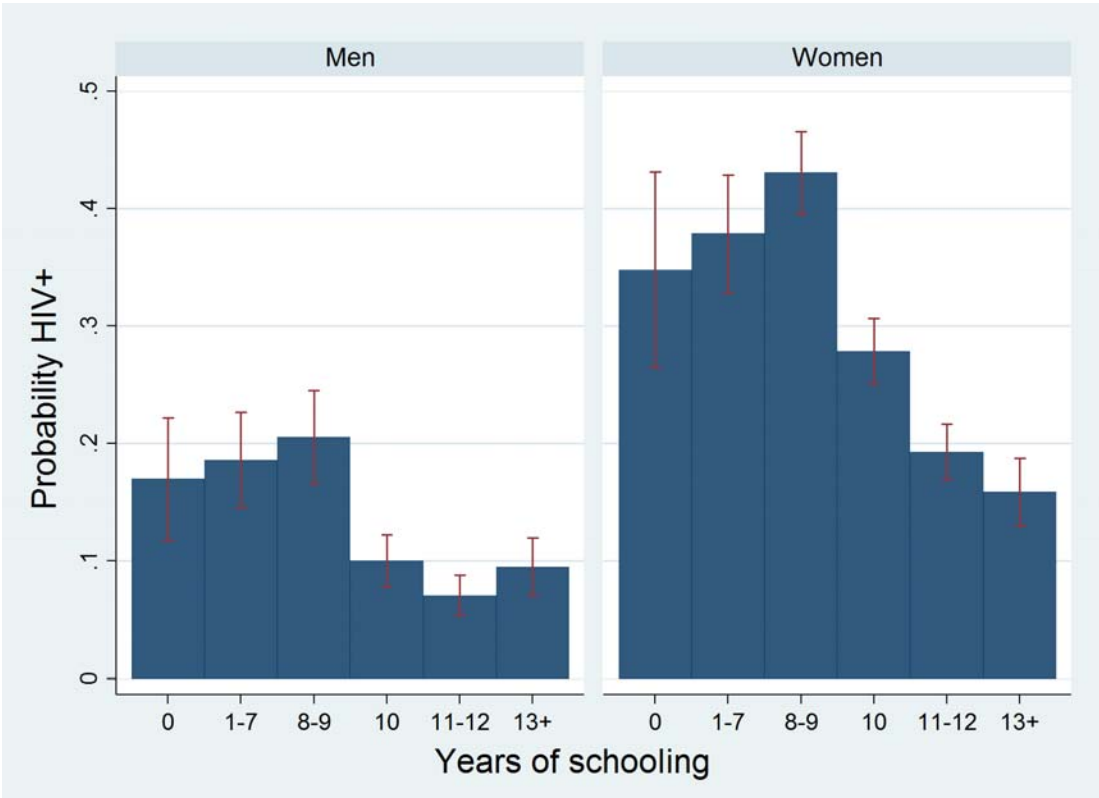
Government of Botswana (1991). National Development Plan - 7, Government printer, Gaborone.

Government of Botswana (1997). National Population Policy, Government printer, Gaborone.

Government of Botswana (2002). Botswana National ART Program.

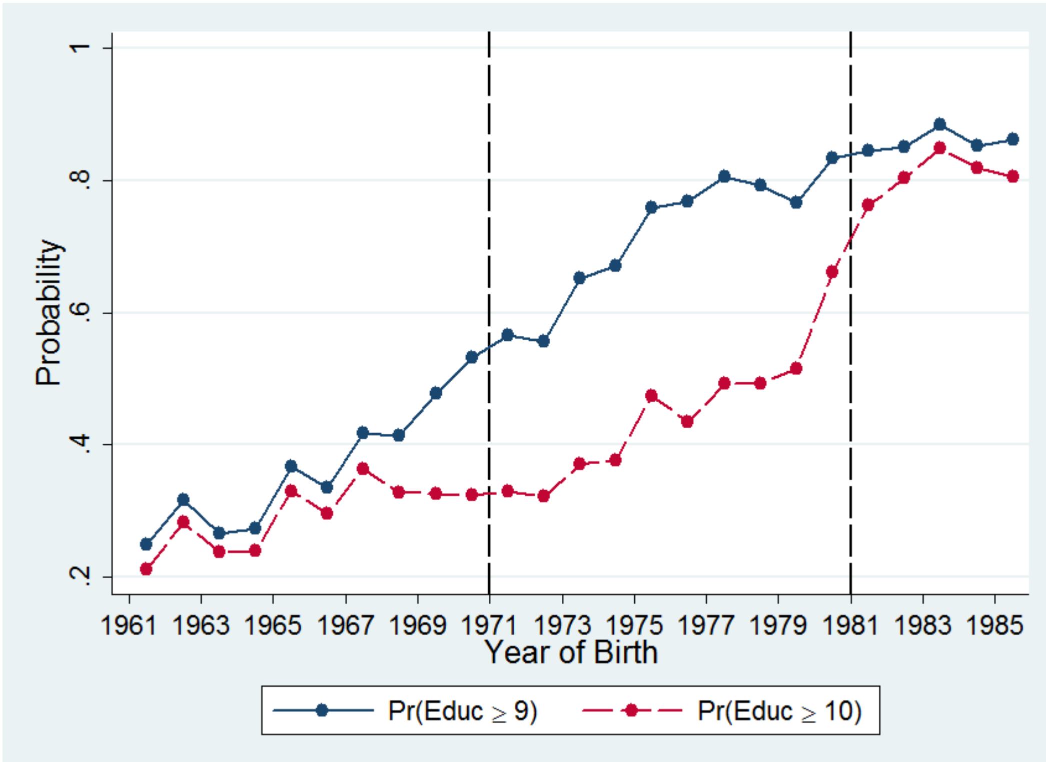
A.4 Supplementary Figures and Tables

Fig. A1: *HIV Prevalence by Years of Schooling in Botswana*



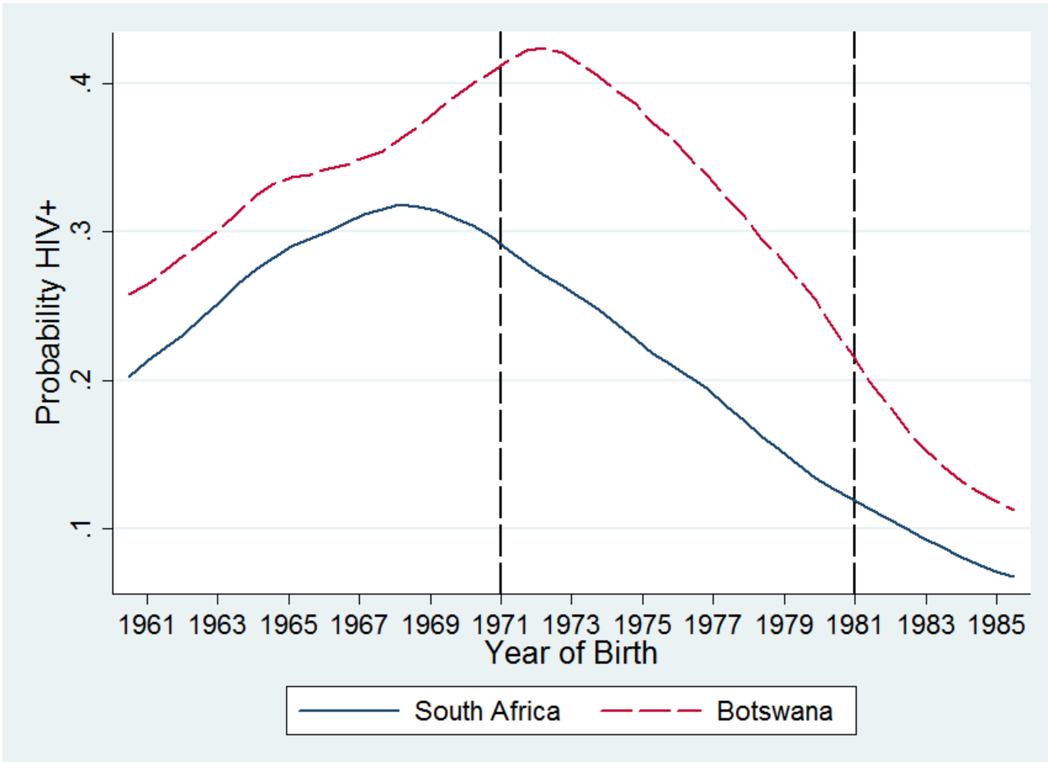
Notes: HIV prevalence by years of schooling completed and gender. Sample includes survey respondents who were citizens of Botswana, at least 20 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. Error bars represent 95% confidence intervals. Survey weights used as provided. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Fig. A2: *Educational Attainment by Birth Cohort in Botswana (Both Reforms)*



Notes: Pr(Educ > 9) is the probability that the respondent has attained at least 9 years of schooling. Pr(Educ > 10) is the probability that the respondent has attained at least 10 years of schooling. Sample includes survey respondents who were citizens of Botswana at least 18 years old at the time of the surveys, born between 1960 and 1985, and had a valid HIV test result. Survey weights used as provided. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Fig. A3: *HIV Prevalence in Botswana and Ethnic Tswana in South Africa*



Notes: Local polynomial smooth plots. Sample includes citizens of Botswana and South Africa born between 1960 and 1985 with at least nine years of schooling (i.e., those most likely to be affected by the policy reform in Botswana). Source: Botswana AIDS Impact Survey II (2004) and III (2008), South African HIV/AIDS Behavioral Risks, Sero-Status, and Mass Media Impact Survey, 2002, and South African National HIV Prevalence, HIV Incidence, Behavior and Communication Survey, 2005.

Table A1: *OLS Regressions: Association Between Years of Schooling and HIV Status.*

Coefficients on Schooling Measures	(1) Female	(2) Female	(3) Male	(4) Male	(5) Both Sexes	(6) Both Sexes	(7) Both Sexes
Predictor							
Years of Schooling	-0.019***		-0.007***		-0.013***		
	(0.002)		(0.002)		(0.001)		
Years of Schooling (0-9)		0.004		0.001		0.003	
		(0.004)		(0.003)		(0.002)	
Years of Schooling (10-13)		-0.045***		-0.022***		-0.036***	
		(0.006)		(0.005)		(0.004)	
At Least 7 Years							-0.003
							(0.029)
At Least 8 Years							0.024
							(0.045)
At Least 9 Years							-0.007
							(0.041)
At Least 10 Years							-0.035*
							(0.018)
At Least 11 Years							-0.038*
							(0.020)
At Least 12 Years							-0.042**
							(0.019)
At Least 13 Years							-0.048***
							(0.013)
Observations	3,965	3,541	3,053	2,658	7,018	6,199	7,018
R-squared	0.111	0.126	0.075	0.089	0.132	0.151	0.141

Notes: Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. All regressions include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models 5 to 7 additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A2: *First Stage Regressions: Effect of the Education Reform on Years of Schooling*

Coefficient on Reform Indicator	**(1)** **Female**	**(2)** **Male**	**(3)** **Both Sexes**	**(4)** **Both Sexes**
Dependent Variable				
Years of Schooling	0.635***	1.005***	0.792***	
	(0.223)	(0.322)	(0.188)	
At Least 7 Years of Schooling				0.026
				(0.016)
At Least 8 Years of Schooling				0.043**
				(0.019)
At Least 9 Years of Schooling				0.042**
				(0.020)
At Least 10 Years of Schooling				0.249***
				(0.024)
At Least 11 Years of Schooling				0.069***
				(0.026)
At Least 12 Years of Schooling				0.082***
				(0.026)
At Least 13 Years of Schooling				0.031
				(0.020)
Observations	3,965	3,053	7,018	7,018
R-squared	0.034	0.033	0.036	-

Notes: All regressions include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Columns 3 and 4 additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights used. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A3.A: *Intention-To-Treat Results: The Effect of the Education Reform on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy (Panel A)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variable	Ever Had Intercourse (1=yes, 0=no)	Age at First Intercourse	First sex ever: Did you use anything to protect yourself (eg, condom)? (1=yes, 0=no)	Ever Married (1=yes, 0=no)	Years of Pre-marital Intercourse	Ever Given Birth (1=yes, 0=no)	How many times have you given birth? (never given birth=0)	Did you have sexual intercourse in the last 12 months? (1=yes, 0=no or never had intercourse)	Last sex: Did you or your partner use a condom, conditional on intercourse in the last 12 months? (1=yes, 0=no)	How old is your most recent partner, conditional on intercourse in the last 12 months? (years)
Coefficient on Reform Indicator										
Female	-0.004	0.592***	0.092***	0.013	-0.688***	-0.114***	-0.181**	-0.008	0.094***	-0.207
	(0.017)	(0.155)	(0.027)	(0.015)	(0.167)	(0.033)	(0.079)	(0.023)	(0.035)	(0.361)
Observations	3,965	3,507	3,582	3,963	3,497	3,644	3,656	3,954	3,041	3,041
R-squared	0.178	0.089	0.078	0.072	0.621	0.198	0.228	0.162	0.024	0.440
Mean Dependent Variable, Pre-Reform	0.964	18.4	0.792	0.064	6.327	0.785	1.378	0.894	0.712	30.3
Male	0.057**	0.073	0.064**	0.010	-0.076	-	-	0.058*	-0.019	-0.510*
	(0.026)	(0.237)	(0.030)	(0.008)	(0.241)	-	-	(0.030)	(0.034)	(0.288)
Observations	3,050	2,360	2,458	3,052	2,357	-	-	3,046	2,144	2,131
R-squared	0.279	0.108	0.047	0.047	0.548	-	-	0.239	0.043	0.44
Mean Dependent Variable, Pre-Reform	0.928	18.6	0.853	0.026	6.683	-	-	0.873	0.822	22.9
Both Sexes	0.022	0.386***	0.081***	0.012	-0.444***	-	-	0.020	0.048*	-0.332
	(0.015)	(0.132)	(0.020)	(0.009)	(0.139)	-	-	(0.018)	(0.025)	(0.244)
Observations	7,015	5,867	6,040	7,015	5,854	-	-	7,000	5,185	5,172
R-squared	0.253	0.099	0.069	0.078	0.592	-	-	0.211	0.046	0.610
Mean Dependent Variable, Pre-Reform	0.948	18.4	0.818	0.048	6.481	-	-	0.885	0.761	27.0

All regressions are OLS models, which include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Years of Premarital Intercourse was defined as Age at First Marriage minus Age at First Intercourse (if married) or Age at Time of Survey minus Age at First Intercourse (if never married). Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A3.B: *Intention-To-Treat Results: The Effect of the Education Reform on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy (Panel B)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variable	Ever had a sexual partner who was 10 years older / younger, conditional on intercourse in the last 12 months? (1=yes, 0=no)	Indicator for 2 or more sexual partners in the last 12 months (1=two or more, 0=one or zero)	Time since first sexual intercourse with most recent partner, conditional on intercourse in the last 12 months (days)	Relationship to most recent partner, conditional on intercourse in the last 12 months? (1=spouse, live-in partner, 0=not living in, other)	Have you ever heard of the virus HIV or an illness called AIDS? (1=yes, 0=no)	Have you ever been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	In the past 12 months have you been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	Indicator for knowledge of at least one HIV prevention strategy (1=yes, 0=no)	Indicator for knowledge of at least two HIV prevention strategies (1=yes, 0=no)	Mentioned condom use as an HIV prevention strategy (1=yes, 0=no)
Coefficient on Reform Indicator										
Female	0.002	0.025	5.8	-0.057	-0.006	0.066**	0.067*	-0.013	-0.014	-0.026
	(0.031)	(0.023)	(84.6)	(0.038)	(0.011)	(0.031)	(0.035)	(0.014)	(0.033)	(0.022)
Observations	3,364	3,658	2,953	3,057	3,874	3,793	3,787	3,791	3,791	3,795
R-squared	0.035	0.016	0.171	0.089	0.011	0.234	0.098	0.020	0.020	0.025
Mean Dependent Variable, Pre-Reform	0.232	0.099	1,285	0.470	0.990	0.683	0.493	0.956	0.706	0.896
Male	-0.027	0.109***	-36.7	0.048	0.005	0.110***	0.069*	0.006	0.004	0.001
	(0.027)	(0.037)	(76.6)	(0.039)	(0.012)	(0.038)	(0.037)	(0.019)	(0.040)	(0.029)
Observations	2,328	2,877	2,086	2,155	2,995	2,922	2,918	2,919	2,919	2,922
R-squared	0.043	0.041	0.147	0.146	0.013	0.249	0.137	0.016	0.018	0.022
Mean Dependent Variable, Pre-Reform	0.114	0.220	894	0.266	0.980	0.551	0.342	0.951	0.678	0.839
Both Sexes	-0.010	0.061***	-11.7	-0.014	-0.001	0.085***	0.067***	-0.005	-0.006	-0.015
	(0.022)	(0.021)	(59.0)	(0.028)	(0.008)	(0.024)	(0.025)	(0.011)	(0.025)	(0.018)
Observations	5,692	6,535	5,039	5,212	6,869	6,715	6,705	6,710	6,710	6,717
R-squared	0.059	0.056	0.195	0.132	0.012	0.28	0.137	0.019	0.022	0.025
Mean Dependent Variable, Pre-Reform	0.180	0.153	1,108	0.379	0.985	0.626	0.427	0.954	0.693	0.871

All regressions are OLS models, which include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The indicator for knowledge of HIV prevention strategies was defined as 1 if respondent could name either at least one (model 8) or two (model 9) out of the six following HIV prevention strategies: condoms, fewer partners, mutually faithful relationship, abstinence, avoid injections with contaminated needles, and avoid blood transfusions. Those responding "don't know" to an HIV knowledge question were accounted for as incorrect. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A3.C: *Intention-To-Treat Results: The Effect of the Education Reform on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy (Panel C)*

Dependent Variable	**(1)** Mentioned mutually faithful relationship as an HIV prevention strategy (1=yes, 0=no)	**(2)** Mentioned abstinence as an HIV prevention strategy (1=yes, 0=no)	**(3)** Can people reduce their chances of getting HIV/AIDS by using a condom correctly every time they have sex? (1=yes, 0=no, don't know)	**(4)** Can people reduce their chances of getting HIV/AIDS by having only one uninfected sex partner who has no other partners? (1=yes, 0=no,	**(5)** Indicator for any misperceptions about HIV (1=yes, 0=no)	**(6)** Do you think it should be acceptable for a woman to obtain male condoms? (1=yes, 0=no or not sure)	**(7)** During the past 4 weeks, have you discussed HIV/AIDS with anyone? (1=yes, 0=no or not sure)	**(8)** During the past 4 weeks, have you discussed HIV/AIDS with your spouse? (1=yes, 0=no)	**(9)** During the past 4 weeks, have you discussed HIV/AIDS with your sexual partner? (1=yes, 0=no)	**(10)** During the past 4 weeks, have you discussed HIV/AIDS with your family member(s)? (1=yes, 0=no)
Coefficient on Reform Indicator										
Female	0.044	0.021	0.007	-0.003	-0.036	0.060***	0.002	-0.002	0.058***	-0.014
	(0.031)	(0.035)	(0.020)	(0.024)	(0.036)	(0.023)	(0.036)	(0.014)	(0.022)	(0.020)
Observations	3,794	3,793	3,791	3,780	3,782	3,873	3,791	3,791	3,791	3,791
R-squared	0.038	0.017	0.015	0.012	0.023	0.037	0.031	0.023	0.020	0.019
Mean Dependent Variable, Pre-Reform	0.230	0.616	0.924	0.876	0.585	0.900	0.483	0.045	0.110	0.091
Male	-0.036	-0.002	0.023	0.033	-0.034	0.074***	0.081**	0.005	0.003	-0.032**
	(0.035)	(0.042)	(0.025)	(0.029)	(0.042)	(0.029)	(0.041)	(0.011)	(0.024)	(0.016)
Observations	2,921	2,921	2,917	2,912	2,915	2,992	2,918	2,918	2,918	2,918
R-squared	0.042	0.018	0.017	0.015	0.019	0.036	0.039	0.02	0.039	0.01
Mean Dependent Variable, Pre-Reform	0.253	0.635	0.898	0.863	0.609	0.868	0.455	0.022	0.075	0.040
Both Sexes	0.010	0.011	0.014	0.012	-0.035	0.066***	0.036	0.001	0.034**	-0.022
	(0.023)	(0.027)	(0.015)	(0.018)	(0.027)	(0.018)	(0.027)	(0.009)	(0.016)	(0.013)
Observations	6,715	6,714	6,708	6,692	6,697	6,865	6,709	6,709	6,709	6,709
R-squared	0.04	0.018	0.016	0.013	0.024	0.04	0.035	0.025	0.028	0.028
Mean Dependent Variable, Pre-Reform	0.240	0.624	0.913	0.870	0.595	0.886	0.471	0.035	0.095	0.069

All regressions are OLS models, which include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The indicator for any misperceptions was defined as 1 if respondent incorrectly answered any of the following four questions: whether HIV spreads via mosquitos, sharing a meal with an HIV+ person, due to witchcraft, and whether a healthy looking person can be HIV+. Those responding "don't know" to an HIV knowledge or HIV misconception question were accounted for as incorrect. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A3.D: *Intention-To-Treat Results: The Effect of the Education Reform on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy (Panel D)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variable	During the past 4 weeks, have you discussed HIV/AIDS with other relative(s)? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a friend? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a co-worker? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a healthcare worker? (1=yes, 0=no)	In the past 4 weeks, have you heard or seen any information about HIV/AIDS? (1=yes, 0=no)	Labor Force Participation (1=yes, 0=no)	Unemployed (1= yes, 0=no, missing if does not participate in labor force)	Indicator for housework (1=worked at home/household/homemaker, 0=other)	Can you read and understand a letter / newspaper / bible? (1=easily, 0=no or with difficulty)	Mentioned newspaper as a source of information about HIV/AIDS in past 4 weeks (1=yes, 0=no)
Coefficient on Reform Indicator										
Female	-0.013	0.009	-0.013	0.014	-0.011	0.108***	0.017	-0.043	0.002	0.041
	(0.016)	(0.028)	(0.017)	(0.016)	(0.035)	(0.033)	(0.043)	(0.030)	(0.023)	(0.030)
Observations	3,791	3,791	3,791	3,791	3,774	3,942	2,458	3,925	3,962	3,778
R-squared	0.020	0.013	0.04	0.019	0.018	0.078	0.128	0.021	0.035	0.015
Mean Dependent Variable, Pre-Reform	0.062	0.157	0.084	0.057	0.649	0.654	0.416	0.261	0.865	0.199
Male	0.014	0.025	0.013	-0.001	0.049	0.048	0.061	-0.033	0.085***	0.060*
	(0.015)	(0.035)	(0.022)	(0.012)	(0.040)	(0.034)	(0.043)	(0.026)	(0.032)	(0.037)
Observations	2,918	2,918	2,918	2,918	2,909	3,037	2,239	3,031	3,051	2,911
R-squared	0.025	0.018	0.058	0.012	0.026	0.156	0.158	0.034	0.029	0.012
Mean Dependent Variable, Pre-Reform	0.023	0.244	0.111	0.031	0.662	0.842	0.219	0.095	0.825	0.214
Both Sexes	-0.001	0.015	-0.002	0.007	0.014	0.082***	0.038	-0.039*	0.037*	0.049**
	(0.011)	(0.022)	(0.013)	(0.010)	(0.026)	(0.024)	(0.030)	(0.021)	(0.019)	(0.023)
Observations	6,709	6,709	6,709	6,709	6,683	6,979	4,697	6,956	7,013	6,689
R-squared	0.023	0.018	0.051	0.023	0.022	0.121	0.154	0.053	0.035	0.016
Mean Dependent Variable, Pre-Reform	0.045	0.195	0.096	0.046	0.655	0.736	0.320	0.189	0.848	0.205

All regressions are OLS models, which include age dummies, a linear term for year of birth, an indicator for survey wave and dummies for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A4.A: *2SLS Results: The Effect of Education on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy (Panel A)*

Dependent Variable	**(1)** Ever Had Intercourse (1=yes, 0=no)	**(2)** Age at First Intercourse	**(3)** First sex ever: Did you use anything to protect yourself (eg, condom)? (1=yes, 0=no)	**(4)** Ever Married (1=yes, 0=no)	**(5)** Years of Pre-marital Intercourse	**(6)** Ever Given Birth (1=yes, 0=no)	**(7)** How many times have you given birth? (never given birth=0)	**(8)** Did you have sexual intercourse in the last 12 months? (1=yes, 0=no or never had intercourse)	**(9)** Last sex: Did you or your partner use a condom, conditional on intercourse in the last 12 months? (1=yes, 0=no)	**(10)** How old is your most recent partner, conditional on intercourse in the last 12 months? (years)
Coefficient on Years of Schooling										
Female	-0.007	0.761***	0.127***	0.021	-0.893***	-0.158***	-0.246**	-0.013	0.142**	-0.311
	(0.027)	(0.261)	(0.047)	(0.025)	(0.304)	(0.057)	(0.107)	(0.035)	(0.068)	(0.538)
Observations	3,965	3,507	3,582	3,963	3,498	3,644	3,656	3,954	3,041	3,041
F-Statistic	8.6	11.9	9.9	8.3	11.6	10.0	10.4	9.1	7.0	7.0
Mean Dependent Variable, Pre-Reform	0.962	18.6	0.828	0.071	6.1	0.728	1.286	0.890	0.759	30.1
Male	0.056*	0.065	0.055*	0.01	-0.084	-	-	0.058*	-0.019	-0.561
	(0.030)	(0.209)	(0.028)	(0.009)	(0.212)	-	-	(0.033)	(0.036)	(0.385)
Observations	3,050	2,360	2,458	3,052	2,357	-	-	3,046	2,144	2,131
F-Statistic	9.6	10.4	11.0	9.6	10.5	-	-	9.2	7.1	6.0
Mean Dependent Variable, Pre-Reform	0.932	18.6	0.863	0.027	6.7	-	-	0.878	0.819	22.8
Both Sexes	0.027	0.421***	0.090***	0.015	-0.493***	-	-	0.025	0.060*	-0.433
	(0.019)	(0.161)	(0.026)	(0.012)	(0.175)	-	-	(0.023)	(0.032)	(0.329)
Observations	7,015	5,867	6,040	7,015	5,855	-	-	7,000	5,185	5,172
F-Statistic	18.1	22.3	20.9	17.8	22.1	-	-	18.3	14.0	13.0
Mean Dependent Variable, Pre-Reform	0.954	18.5	0.840	0.051	6.4	-	-	0.890	0.776	26.9

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Years of Premarital Intercourse was defined as Age at First Marriage minus Age at First Intercourse (if married) or Age at Time of Survey minus Age at First Intercourse (if never married). Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A4.B: *2SLS Results: The Effect of Education on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy (Panel B)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variable	Ever had a sexual partner who was 10 years older / younger, conditional on intercourse in the last 12 months? (1=yes, 0=no)	Indicator for 2 or more sexual partners in the last 12 months (1=two or more, 0=one or zero)	Time since first sexual intercourse with most recent partner, conditional on intercourse in the last 12 months (days)	Relationship to your most recent partner, conditional on intercourse in the last 12 months? (1=spouse, live-in partner, 0=not living in, other)	Have you ever heard of the virus HIV or an illness called AIDS? (1=yes, 0=no)	Have you ever been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	In the past 12 months have you been tested for HIV, the virus that causes AIDS? (1=yes, 0=no)	Indicator for knowledge of at least one HIV prevention strategy (1=yes, 0=no)	Indicator for knowledge of at least two HIV prevention strategies (1=yes, 0=no)	Mentioned condom use as an HIV prevention strategy (1=yes, 0=no)
Coefficient on Years of Schooling										
Female	0.003	0.044	8.1	-0.087	-0.009	0.110*	0.111	-0.021	-0.023	-0.044
	(0.042)	(0.045)	(118.5)	(0.063)	(0.017)	(0.061)	(0.068)	(0.020)	(0.056)	(0.041)
Observations	3,364	3,658	2,953	3,057	3,874	3,793	3,787	3,791	3,791	3,795
F-Statistic	9.7	6.3	8.0	6.8	8.3	7.7	7.6	7.9	7.9	7.8
Mean Dependent Variable, Pre-Reform	0.232	0.115	1,289	0.441	0.987	0.720	0.531	0.949	0.698	0.881
Male	-0.026	0.111**	-39.5	0.049	0.005	0.120**	0.075*	0.007	0.005	0.002
	(0.026)	(0.051)	(82.0)	(0.045)	(0.012)	(0.052)	(0.044)	(0.020)	(0.043)	(0.031)
Observations	2,328	2,877	2,086	2,155	2,995	2,922	2,918	2,919	2,919	2,922
F-Statistic	8.3	8.5	6.0	6.7	9.4	7.9	7.8	7.9	7.9	8.1
Mean Dependent Variable, Pre-Reform	0.112	0.232	888	0.273	0.980	0.573	0.355	0.952	0.678	0.840
Both Sexes	-0.011	0.082**	-14.6	-0.018	-0.001	0.115***	0.092**	-0.006	-0.009	-0.02
	(0.025)	(0.034)	(72.8)	(0.035)	(0.010)	(0.040)	(0.039)	(0.016)	(0.035)	(0.025)
Observations	5,692	6,535	5,039	5,212	6,869	6,715	6,705	6,710	6,710	6,717
F-Statistic	18.2	14.9	13.9	13.5	17.8	15.6	15.4	15.8	15.8	15.8
Mean Dependent Variable, Pre-Reform	0.178	0.173	1,104	0.375	0.985	0.657	0.452	0.952	0.691	0.866

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. The indicator for knowledge of HIV prevention strategies was defined as 1 if respondent could name either at least one (model 8) or two (model 9) out of the six following HIV prevention strategies: condoms, fewer partners, mutually faithful relationship, abstinence, avoid injections with contaminated needles, and avoid blood transfusions. Those responding "don't know" to an HIV knowledge question were accounted for as incorrect. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A4.C: *2SLS Results: The Effect of Education on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy (Panel C)*

Dependent Variable	**(1)** Mentioned mutually faithful relationship as an HIV prevention strategy (1=yes, 0=no)	**(2)** Mentioned abstinence as an HIV prevention strategy (1=yes, 0=no)	**(3)** Can people reduce their chances of getting HIV/AIDS by using a condom correctly every time they have sex? (1=yes, 0=no, don't know)	**(4)** Can people reduce their chances of getting HIV/AIDS by having only one uninfected sex partner who has no other partners? (1=yes, 0=no,	**(5)** Indicator for any misperceptions about HIV (1=yes, 0=no)	**(6)** Do you think it should be acceptable for a woman to obtain male condoms? (1=yes, 0=no or not sure)	**(7)** During the past 4 weeks, have you discussed HIV/AIDS with anyone? (1=yes, 0=no or not sure)	**(8)** During the past 4 weeks, have you discussed HIV/AIDS with your spouse? (1=yes, 0=no)	**(9)** During the past 4 weeks, have you discussed HIV/AIDS with your sexual partner? (1=yes, 0=no)	**(10)** During the past 4 weeks, have you discussed HIV/AIDS with your family member(s)? (1=yes, 0=no)
Coefficient on Years of Schooling										
Female	0.073	0.034	0.012	-0.005	-0.062	0.092**	0.003	-0.003	0.094**	-0.022
	(0.054)	(0.057)	(0.032)	(0.040)	(0.059)	(0.040)	(0.060)	(0.023)	(0.047)	(0.035)
Observations	3,794	3,793	3,791	3,780	3,782	3,873	3,791	3,791	3,792	3,792
F-Statistic	7.7	7.7	7.8	7.7	7.2	8.7	7.7	7.7	7.7	7.7
Mean Dependent Variable, Pre-Reform	0.254	0.628	0.928	0.874	0.564	0.933	0.484	0.044	0.142	0.083
Male	-0.038	-0.002	0.026	0.038	-0.038	0.072**	0.087*	0.006	0.003	-0.035
	(0.043)	(0.045)	(0.027)	(0.034)	(0.043)	(0.030)	(0.046)	(0.012)	(0.026)	(0.022)
Observations	2,921	2,921	2,917	2,912	2,915	2,992	2,918	2,918	2,918	2,918
F-Statistic	8.0	8.1	7.7	7.2	7.6	9.6	8.2	8.2	8.2	8.2
Mean Dependent Variable, Pre-Reform	0.246	0.634	0.902	0.869	0.603	0.875	0.471	0.023	0.075	0.034
Both Sexes	0.014	0.015	0.019	0.017	-0.049	0.082***	0.048	0.002	0.045*	-0.029
	(0.031)	(0.035)	(0.021)	(0.025)	(0.036)	(0.025)	(0.035)	(0.012)	(0.024)	(0.020)
Observations	6,715	6,714	6,708	6,692	6,697	6,865	6,709	6,709	6,710	6,710
F-Statistic	15.8	15.7	15.4	14.8	14.7	18.3	15.8	15.8	15.8	15.8
Mean Dependent Variable, Pre-Reform	0.244	0.628	0.918	0.875	0.582	0.905	0.484	0.036	0.107	0.061

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. The indicator for any misperceptions was defined as 1 if respondent incorrectly answered any of the following four questions: whether HIV spreads via mosquitos, sharing a meal with an HIV+ person, due to witchcraft, and whether a healthy looking person can be HIV+. Those responding "don't know" to an HIV knowledge or HIV misconception question were accounted for as incorrect. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A4.D: *2SLS Results: The Effect of Education on Sexual Intercourse, Marriage, Current Sexual and HIV Testing Behavior, HIV Knowledge and Attitudes, Labor Market Outcomes, and Literacy (Panel D)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variable	During the past 4 weeks, have you discussed HIV/AIDS with other relative(s)? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a friend? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a co-worker? (1=yes, 0=no)	During the past 4 weeks, have you discussed HIV/AIDS with a healthcare worker? (1=yes, 0=no)	In the past 4 weeks, have you heard or seen any information about HIV/AIDS? (1=yes, 0=no)	Labor Force Participation (1=yes, 0=no)	Unemployed (1= yes, 0=no, missing if does not participate in labor force)	Indicator for housework (1=worked at home/househol d/homemaker, 0=other)	Can you read and understand a letter / newspaper / bible? (1=easily, 0=no or with difficulty)	Mentioned newspaper as a source of information about HIV/AIDS in past 4 weeks (1=yes, 0=no)
Coefficient on Years of Schooling										
Female	-0.022	0.015	-0.022	0.023	-0.019	0.172**	0.038	-0.070	0.003	0.069
	(0.028)	(0.046)	(0.030)	(0.027)	(0.060)	(0.076)	(0.100)	(0.049)	(0.036)	(0.049)
Observations	3,791	3,793	3,791	3,791	3,774	3,942	2,458	3,925	3,962	3,778
F-Statistic	7.7	7.7	7.7	7.7	7.6	8.4	2.8	8.2	8.7	7.6
Mean Dependent Variable, Pre-Reform	0.055	0.163	0.077	0.065	0.643	0.706	0.433	0.239	0.866	0.222
Male	0.015	0.026	0.014	-0.001	0.052	0.048	0.069	-0.033	0.084***	0.064*
	(0.017)	(0.037)	(0.023)	(0.012)	(0.041)	(0.039)	(0.056)	(0.027)	(0.023)	(0.038)
Observations	2,918	2,920	2,918	2,918	2,909	3,037	2,239	3,031	3,051	2,911
F-Statistic	8.2	8.0	8.2	8.2	8.2	9.3	5.4	9.3	9.7	8.3
Mean Dependent Variable, Pre-Reform	0.026	0.249	0.114	0.031	0.671	0.846	0.235	0.092	0.831	0.224
Both Sexes	-0.002	0.021	-0.003	0.01	0.019	0.105***	0.058	-0.050*	0.046**	0.066**
	(0.015)	(0.029)	(0.018)	(0.014)	(0.034)	(0.039)	(0.051)	(0.027)	(0.019)	(0.031)
Observations	6,709	6,713	6,709	6,709	6,683	6979	4,697	6,956	7,013	6,689
F-Statistic	15.8	15.7	15.8	15.8	15.7	17.7	8.2	17.5	18.4	15.9
Mean Dependent Variable, Pre-Reform	0.045	0.201	0.095	0.049	0.660	0.757	0.340	0.178	0.857	0.223

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. Models for both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, born in or after 1975, and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A5.A: *Sensitivity Analyses in the Pooled Sample: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy (Panel A)*

Dependent Variable: HIV Positive ***Sample: Both Sexes*** ***Model: 2SLS***	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Coefficient on endogenous variable										
Years of Schooling	-0.081***	-0.070**	-0.067**	-0.073**	-0.069**	-0.061*	-0.095*	-0.112	-0.078**	-0.109**
	(0.031)	(0.035)	(0.033)	(0.030)	(0.027)	(0.033)	(0.054)	(0.100)	(0.035)	(0.049)
Basic Covariates										
i.Age, i.Age#i.Sex	✓	-	-	✓	✓	✓	✓	✓	✓	✓
c.YOB, c.YOB#i.Sex	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.Survey_Year, i.Survey_Year#i.Sex	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.DistrictofBirth, i.Sex, i.DistrictofBirth#i.Sex	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.District_Edu9##i.Sex, i.Reform_Indicator##i.Sex	-	-	-	-	-	-	-	-	-	-
Additional Controls										
c.Age##i.Sex, c.Age2##i.Sex	-	✓	-	-	-	-	-	-	-	-
Cubic Spline (f(Age)*Sex): c.Age#i.Sex, c.Age2##i.Sex, c.Age3##i.Sex, c.Agek20##i.Sex, c.Agek25##i.Sex, c.Agek30##i.Sex	-	-	✓	-	-	-	-	-	-	-
c.YOB2##i.Sex	-	-	-	✓	-	-	-	-	-	-
c.YOB##i.Survey_Year##i.Sex	✓	✓	✓	-	-	✓	✓	✓	✓	-
c.YOB##i.Survey_Year##i.Sex, c.YOB2##i.Survey_Year##i.Sex	-	-	-	-	✓	-	-	-	-	-
c.YOB#i.Reform_Indicator, c.YOB#i.Reform_Indicator#i.Sex	-	-	-	-	-	✓	-	-	-	-
Narrower Birth Cohort Windows										
YOB > 1974	✓	✓	✓	✓	✓	✓	-	-	✓	✓
1975 < YOB < 1987	-	-	-	-	-	-	✓	-	-	-
1976 < YOB < 1986	-	-	-	-	-	-	-	✓	-	-
Sample Weights	-	-	-	-	-	-	-	-	✓	-
Imputed HIV Estimates	-	-	-	-	-	-	-	-	-	✓
Excluding 1980 Birth Cohort	-	-	-	-	-	-	-	-	-	-
Subsample, Less Than Nine Years of Schooling	-	-	-	-	-	-	-	-	-	-
Subsample, At Least Nine Years of Schooling	-	-	-	-	-	-	-	-	-	-
Earlier Cohorts Included, 1903 < YOB < 1990	-	-	-	-	-	-	-	-	-	-
Observations	7,018	7,018	7,018	7,018	7,018	7,018	5,712	4,987	7,018	10,039
F-Statistic	17.1	12.5	14.8	17.9	0.9	13.3	5.9	1.9	13.4	n/a

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. Model 3 controls for a third order spline in age (with knots at 20, 25, and 30) interacted with sex (Durrleman et al. 1989). The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A5.B: *Sensitivity Analyses in the Pooled Sample: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy (Panel B)*

Dependent Variable: HIV Positive ***Sample: Both Sexes*** ***Model: 2SLS***	**(11)¥**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**	**(17)#**	**(18)**	**(19)**	**(20)**
Coefficient on endogenous variable										
Years of Schooling	-0.091**	-0.123*	-0.052**	0.140	-0.146***	-0.177***	-0.052***	-0.063**	-0.099**	-0.081***
	(0.044)	(0.074)	(0.025)	(0.093)	(0.053)	(0.057)	(0.014)	(0.031)	(0.041)	(0.031)
Basic Covariates										
i.Age, i.Age#i.Sex	✓	✓	✓	✓	✓	✓	✓	-	-	✓
c.YOB, c.YOB#i.Sex	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
i.Survey_Year, i.Survey_Year#i.Sex	✓	✓	✓	✓	✓	✓	✓	✓	-	✓
i.DistrictofBirth, i.Sex, i.DistrictofBirth##i.Sex	✓	-	✓	✓	✓	✓	✓	✓	✓	✓
i.District_Edu9##i.Sex, i.Reform_Indicator##i.Sex	-	✓	-	-	-	-	-	-	-	-
Additional Controls										
c.Age##i.Sex, c.Age2##i.Sex	-	-	-	-	-	-	-	-	-	-
Cubic Spline (f(Age)*Sex): c.Age#i.Sex, c.Age2##i.Sex, c.Age3##i.Sex, c.Agek20##i.Sex, c.Agek25##i.Sex, c.Agek30##i.Sex	-	-	-	-	-	-	-	-	-	-
c.YOB2##i.Sex	-	-	-	-	-	-	-	-	-	-
c.YOB##i.Survey_Year##i.Sex	-	-	-	-	-	-	-	-	-	-
c.YOB##i.Survey_Year##i.Sex, c.YOB2##i.Survey_Year##i.Sex	-	-	-	-	-	-	-	-	-	-
c.YOB#i.Reform_Indicator,	-	-	-	-	-	-	-	-	-	-
c.YOB#i.Reform_Indicator#i.Sex	-	-	-	-	-	-	-	-	-	-
Narrower Birth Cohort Windows										
YOB > 1974	✓	✓	✓	✓	✓	-	✓	✓	✓	✓
1975 < YOB < 1987	-	-	-	-	-	-	-	-	-	-
1976 < YOB < 1986	-	-	-	-	-	-	-	-	-	-
Sample Weights	-	-	-	-	-	-	-	-	-	-
Imputed HIV Estimates	✓	-	-	-	-	-	-	-	-	-
Excluding 1980 Birth Cohort	-	-	✓	-	-	-	-	-	-	-
Subsample, Less Than Nine Years of Schooling	-	-	-	✓	-	-	-	-	-	-
Subsample, At Least Nine Years of Schooling	-	-	-	-	✓	-	-	-	-	-
Earlier Cohorts Included, 1903 < YOB < 1990	-	-	-	-	-	✓	-	-	-	-
Observations	8,281	7,018	6,460	1,175	5,843	14,674	7,018	7,018	7,018	7,018
F-Statistic	n/a	0.2	23.0	0.1	16.0	12.3	n/a	15	12	18

All regressions are 2SLS models, except models 14 and 15 which are OLS models (ITT), and model 17 which is a probit model using Stata's ivprobit command. Exposure to the reform was used as an instrument for years of schooling. The instrument was a reform indicator defined as = 1 if YOB > 1980, except in model 12 where the instrument was the interation ReformIndicator*DistrictEdu9. DistrictEdu9 was defined as an indicator for a high vs. low proportion of ninth grade completion by district of birth pre-reform using the Botswana Labor Force Surveys 1995/96 and 2005/06. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. ¥ Uses additional covariates, Age at First Intercourse and Ever Married, to impute HIV status. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A5.C: *Sensitivity Analyses in the Female Subsample: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy (Panel C)*

Dependent Variable: HIV Positive ***Sample: Female*** ***Model: 2SLS***	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Coefficient on endogenous variable										
Years of Schooling	-0.116**	-0.093	-0.115	-0.140*	-0.123*	-0.112	-0.279	-0.339	-0.113*	-0.141
	(0.058)	(0.065)	(0.081)	(0.079)	(0.068)	(0.084)	(0.303)	(0.675)	(0.060)	(0.080)
Basic Covariates										
i.Age	✓	-	-	✓	✓	✓	✓	✓	✓	✓
c.YOB	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.Survey_Year	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.DistrictofBirth	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.District_Edu9, i.Reform_Indicator	-	-	-	-	-	-	-	-	-	-
Additional Controls										
c.Age, c.Age2	-	✓	-	-	-	-	-	-	-	-
Cubic Spline (f(Age)): c.Age, c.Age2, c.Age3, c.Agek20, c.Agek25, c.Agek30	-	-	✓	-	-	-	-	-	-	-
c.YOB2	-	-	-	✓	-	-	-	-	-	-
c.YOB##i.Survey_Year	✓	✓	✓	-	-	✓	✓	✓	✓	-
c.YOB##i.Survey_Year, c.YOB2##i.Survey_Year	-	-	-	-	✓	-	-	-	-	-
c.YOB#i.Reform_Indicator	-	-	-	-	-	✓	-	-	-	-
Narrower Birth Cohort Windows										
YOB > 1974	✓	✓	✓	✓	✓	✓	-	-	✓	✓
1975 < YOB < 1987	-	-	-	-	-	-	✓	-	-	-
1976 < YOB < 1986	-	-	-	-	-	-	-	✓	-	-
Sample Weights	-	-	-	-	-	-	-	-	✓	-
Imputed HIV Estimates	-	-	-	-	-	-	-	-	-	✓
Excluding 1980 Birth Cohort	-	-	-	-	-	-	-	-	-	-
Subsample, Less Than Nine Years of Schooling	-	-	-	-	-	-	-	-	-	-
Subsample, At Least Nine Years of Schooling	-	-	-	-	-	-	-	-	-	-
Earlier Cohorts Included, 1903 < YOB < 1990	-	-	-	-	-	-	-	-	-	-
Observations	3,965	3,965	3,965	3,965	3,965	3,965	3,240	2,846	3,965	5,549
F-Statistic	8.3	6.0	4.6	5.3	0.2	3.9	0.9	0.2	7.7	n/a

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. Model 3 controls for a third order spline in age (with knots at 20, 25, and 30) (Durrleman et al. 1989). The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A5.D: *Sensitivity Analyses in the Female Subsample: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy (Panel D)*

Dependent Variable: HIV Positive ***Sample: Female*** ***Model: 2SLS***	**(11)**[¥]	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**	**(17)**[#]	**(18)**	**(19)**	**(20)**
Coefficient on endogenous variable										
Years of Schooling	-0.130*	-0.132**	-0.062*	0.048	-0.112**	-0.308**	-0.066***	-0.081	-0.121*	-0.116**
	(0.070)	(0.060)	(0.038)	(0.128)	(0.046)	(0.135)	(0.022)	(0.056)	(0.072)	(0.058)
Basic Covariates										
i.Age	✓	✓	✓	✓	✓	✓	✓	-	-	✓
c.YOB	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
i.Survey_Year	✓	✓	✓	✓	✓	✓	✓	✓	-	✓
i.DistrictofBirth	✓	-	✓	✓	✓	✓	✓	✓	✓	✓
i.District_Edu9, i.Reform_Indicator	-	✓	-	-	-	-	-	-	-	-
Additional Controls										
c.Age, c.Age2	-	-	-	-	-	-	-	-	-	-
Cubic Spline (f(Age)): c.Age, c.Age2, c.Age3, c.Agek20, c.Agek25, c.Agek30	-	-	-	-	-	-	-	-	-	-
c.YOB2	-	-	-	-	-	-	-	-	-	-
c.YOB##i.Survey_Year	-	-	-	-	-	-	-	-	-	-
c.YOB##i.Survey_Year, c.YOB2##i.Survey_Year	-	-	-	-	-	-	-	-	-	-
c.YOB#i.Reform_Indicator	-	-	-	-	-	-	-	-	-	-
Narrower Birth Cohort Windows										
YOB > 1974	✓	✓	✓	✓	✓	-	✓	✓	✓	✓
1975 < YOB < 1987	-	-	-	-	-	-	-	-	-	-
1976 < YOB < 1986	-	-	-	-	-	-	-	-	-	-
Sample Weights	-	-	-	-	-	-	-	-	-	-
Imputed HIV Estimates	✓	-	-	-	-	-	-	-	-	-
Excluding 1980 Birth Cohort	-	-	✓	-	-	-	-	-	-	-
Subsample, Less Than Nine Years of Schooling	-	-	-	✓	-	-	-	-	-	-
Subsample, At Least Nine Years of Schooling	-	-	-	-	✓	-	-	-	-	-
Earlier Cohorts Included, 1903 < YOB < 1990	-	-	-	-	-	✓	-	-	-	-
Observations	4,866	3,965	3,650	565	3,400	8,456	3,965	3,965	3,965	3,965
F-Statistic	n/a	9.3	8.4	1.5	22.5	5.6	n/a	7.4	5.8	8.6

All regressions are 2SLS models, except model 17, which is a probit model, using Stata's ivprobit command. Exposure to the reform was used as an instrument for years of schooling. The instrument was defined as = 1 if YOB > 1980, except in model 13 where the instrument was defined as 1 if YOB > 1979. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ¥ Uses additional covariates, Age at First Intercourse and Ever Married, to impute HIV status. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A5.E: *Sensitivity Analyses in the Male Subsample: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy (Panel E)*

Dependent Variable: HIV Positive ***Sample: Male*** ***Model: 2SLS***	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coefficient on endogenous variable										
Years of Schooling	-0.050*	-0.05	-0.038	-0.035	-0.033	-0.031	-0.029	-0.034	-0.043	-0.056
	(0.029)	(0.035)	(0.026)	(0.024)	(0.023)	(0.027)	(0.033)	(0.052)	(0.036)	(0.052)
Basic Covariates										
i.Age	✓	-	-	✓	✓	✓	✓	✓	✓	✓
c.YOB	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.Survey_Year	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.DistrictofBirth	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
i.District_Edu9, i.Reform_Indicator	-	-	-	-	-	-	-	-	-	-
Additional Controls										
c.Age, c.Age2	-	✓	-	-	-	-	-	-	-	-
Cubic Spline (f(Age)): c.Age, c.Age2, c.Age3, c.Agek20, c.Agek25, c.Agek30	-	-	✓	-	-	-	-	-	-	-
c.YOB2	-	-	-	✓	-	-	-	-	-	-
c.YOB##i.Survey_Year	✓	✓	✓	-	-	✓	✓	✓	✓	-
c.YOB##i.Survey_Year, c.YOB2##i.Survey_Year	-	-	-	-	✓	-	-	-	-	-
c.YOB#i.Reform_Indicator	-	-	-	-	-	✓	-	-	-	-
Narrower Birth Cohort Windows										
YOB > 1974	✓	✓	✓	✓	✓	✓	-	-	✓	✓
1975 < YOB < 1987	-	-	-	-	-	-	✓	-	-	-
1976 < YOB < 1986	-	-	-	-	-	-	-	✓	-	-
Sample Weights	-	-	-	-	-	-	-	-	✓	-
Imputed HIV Estimates	-	-	-	-	-	-	-	-	-	✓
Excluding 1980 Birth Cohort	-	-	-	-	-	-	-	-	-	-
Subsample, Less Than Nine Years of Schooling	-	-	-	-	-	-	-	-	-	-
Subsample, At Least Nine Years of Schooling	-	-	-	-	-	-	-	-	-	-
Earlier Cohorts Included, 1903 < YOB < 1990	-	-	-	-	-	-	-	-	-	-
Observations	3,053	3,053	3,053	3,053	3,053	3,053	2,472	2,141	3,053	4,490
F-Statistic	8.8	6.6	10.6	13.1	0.7	10.0	6.0	2.1	5.8	n/a

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. Model 3 controls for a third order spline in age (with knots at 20, 25, and 30) (Durrleman et al. 1989). The instrument was defined as = 1 if YOB > 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A5.F: *Sensitivity Analyses in the Male Subsample: Assessing the Robustness of the 2SLS Results to: Additional Controls, Sample Restrictions, Sample Weights, Imputed HIV Estimates, and a 2SLS Difference-in-Differences Strategy (Panel F)*

Dependent Variable: HIV Positive ***Sample: Male*** ***Model: 2SLS***	**(11)[¥]**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**	**(17)[#]**	**(18)**	**(19)**	**(20)**
Coefficient on endogenous variable										
Years of Schooling	-0.053	-0.061*	-0.040	-0.006	-0.381	-0.068*	-0.040**	-0.047	-0.079*	-0.050*
	-0.047	(0.031)	(0.029)	(0.081)	(0.503)	(0.039)	(0.016)	(0.032)	(0.043)	(0.029)
Basic Covariates										
i.Age	✓	✓	✓	✓	✓	✓	✓	-	-	✓
c.YOB	✓	✓	✓	✓	✓	✓	✓	✓	✓	-
i.Survey_Year	✓	✓	✓	✓	✓	✓	✓	✓	-	✓
i.DistrictofBirth	✓	-	✓	✓	✓	✓	✓	✓	✓	✓
i.District_Edu9, i.Reform_Indicator	-	✓	-	-	-	-	-	-	-	-
Additional Controls										
c.Age, c.Age2	-	-	-	-	-	-	-	-	-	-
Cubic Spline (f(Age)): c.Age, c.Age2, c.Age3, c.Agek20, c.Agek25, c.Agek30	-	-	-	-	-	-	-	-	-	-
c.YOB2	-	-	-	-	-	-	-	-	-	-
c.YOB##i.Survey_Year	-	-	-	-	-	-	-	-	-	-
c.YOB##i.Survey_Year, c.YOB2##i.Survey_Year	-	-	-	-	-	-	-	-	-	-
c.YOB#i.Reform_Indicator	-	-	-	-	-	-	-	-	-	-
Narrower Birth Cohort Windows										
YOB > 1974	✓	✓	✓	✓	✓	-	✓	✓	✓	✓
1975 < YOB < 1987	-	-	-	-	-	-	-	-	-	-
1976 < YOB < 1986	-	-	-	-	-	-	-	-	-	-
Sample Weights	-	-	-	-	-	-	-	-	-	-
Imputed HIV Estimates	✓	-	-	-	-	-	-	-	-	-
Excluding 1980 Birth Cohort	-	-	✓	-	-	-	-	-	-	-
Subsample, Less Than Nine Years of Schooling	-	-	-	✓	-	-	-	-	-	-
Subsample, At Least Nine Years of Schooling	-	-	-	-	✓	-	-	-	-	-
Earlier Cohorts Included, 1903 < YOB < 1990	-	-	-	-	-	✓	-	-	-	-
Observations	3,415	3,053	2,810	610	2,443	6,218	3,053	3,053	3,053	3,053
F-Statistic	n/a	10.1	8.5	2.2	0.6	6.7	n/a	7.6	6.0	9.5

All regressions are 2SLS models, except model 17, which is a probit model, using Stata's ivprobit command. Exposure to the reform was used as an instrument for years of schooling. The instrument was defined as = 1 if YOB > 1980, except in model 13 where the instrument was defined as 1 if YOB > 1979. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ¥ Uses additional covariates, Age at First Intercourse and Ever Married, to impute HIV status. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys, and had a valid HIV test result. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A6: *Additional Policy Reform: First Stage, Reduced Form, and 2SLS Results*

	(1)	(2)	(3)
Model	**First Stage**	**Intention-to-treat**	**2SLS**
Dependent Variable	**Years of Schooling**	**HIV-Positive**	**HIV-Positive**
Coefficient on Reform Indicator			
Female	0.261	-0.111***	-0.425*
	(0.165)	(0.021)	(0.249)
Observations	8,456	8,456	8,456
R-squared	0.291	0.081	-
Probability HIV-positive, Reform Period	-	0.467	0.373
F-Statistic	-	-	2.982
Male	0.533**	-0.027	-0.051
	(0.235)	(0.021)	(0.044)
Observations	6,218	6,218	6,218
R-squared	0.220	0.106	-
Probability HIV-positive, Reform Period	-	0.278	0.251
F-Statistic	-	-	4.764
Both Sexes	0.375***	-0.076***	-0.202**
	(0.138)	(0.015)	(0.079)
Observations	14,674	14,674	14,674
R-squared	0.256	0.097	-
Probability HIV-positive, Reform Period	-	0.385	0.312
F-Statistic	-	-	7.735

Regressions 1 to 2 are OLS models. Regression 3 is a 2SLS model, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave and indicators for district of birth. Regressions for the subsample with both sexes additionally control for age*sex, districtofbirth*sex, yearofbirth*sex and surveywave*sex interactions. The instrument was defined as = 1 if YOB between 1971 and 1980. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Botswana, at least 18 years old at the time of the surveys and had a valid HIV test result. No weights were used. Source: Botswana AIDS Impact Survey II (2004) and III (2008).

Table A7: *Falsification Test: First Stage, Reduced Form, and 2SLS Results for Ethnic Tswana South Africans*

Model	**First Stage**	**OLS**	**Intention-to-Treat**	**2SLS**
Dependent Variable	**Schooling (years)**	**HIVpos**	**HIVpos**	**HIVpos**
Predictors				
ReformIndicator	0.157		-0.040	
	(0.746)		(0.080)	
YearsofSchooling		-0.010*		-0.061
		(0.006)		(0.075)
Observations	444	444	444	444
R-squared	0.054	0.129	0.122	-

Regressions 1 to 3 are OLS models. Regression 4 is a 2SLS model, in which exposure to the reform was used as an instrument for years of schooling. All models included the following controls: single-year age indicators, a linear term for year of birth, an indicator for survey wave, and age*sex, yearofbirth*sex, and surveywave*sex interactions. The instrument was defined as = 1 if YOB > 1980. DistrictofBirth was not available in the South African National HIV Survey of 2005 and not included as a covariate. F-statistic: 0.04. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of South Africa, at least 18 years old at the time of the surveys, born after 1975, and had a valid HIV test result. No weights were used. Source: South African HIV/AIDS Behavioral Risks, Sero-Status, and Mass Media Impact Survey, 2002 and South African National HIV Prevalence, HIV Incidence, Behavior and Communication Survey, 2005.

Table A8: *Alternative Hypotheses Possibly Affecting Post-1980 Birth Cohorts*

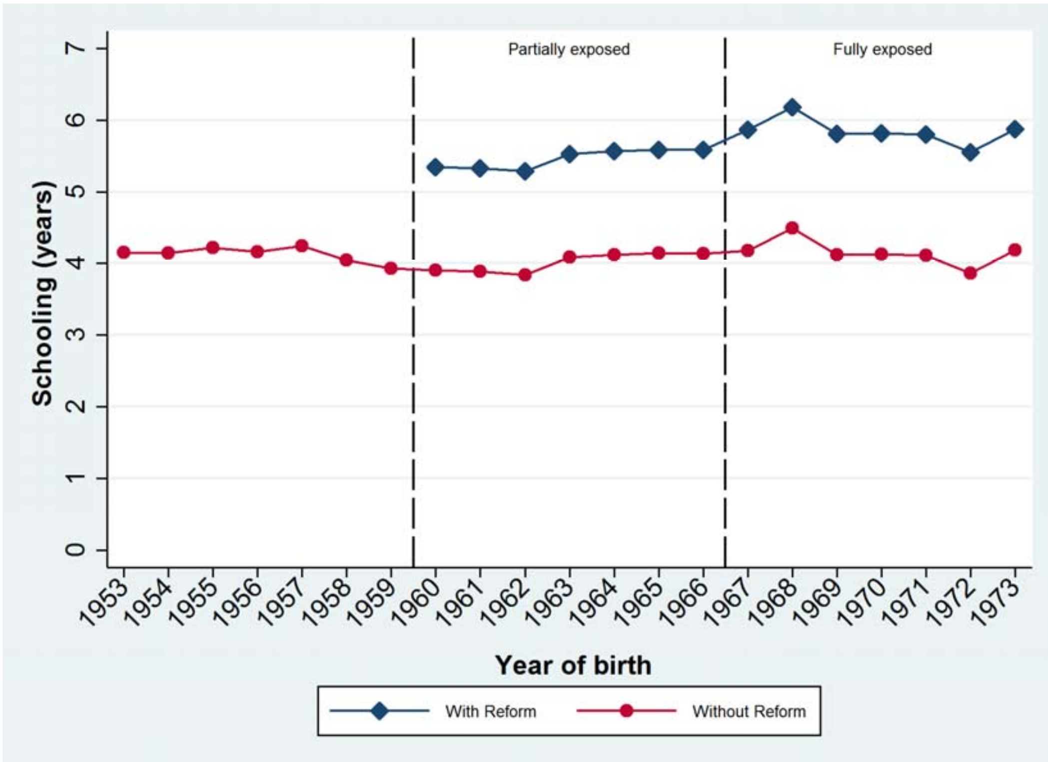
Alternative hypothesis	Policy (year) and description	Notes
Curriculum changes	• Revised National Policy of Education (1994): increase the vocational orientation of academic subjects, increase the number of practical subjects, emphasize professional skills, relate the curriculum to the professional environment, and increase career guidance.	Not implemented nationally; gradual implementation over at least five years.
National ART program.	• National ART Program (2002): ART became available free of charge through public health services. Aimed to deliver care with a high level of clinical monitoring and a low tolerance of adverse events, as in high-resource settings.	Rolled out later and/or did not affect specific birth cohorts.
Home based care, Prevention of Mother to Child Transmission, HIV testing policies	• Community Home Based Care (1995): ensure continuity of comprehensive care services and social support to AIDS patients. • Prevention of Mother to Child Transmission (1999): available in all public health facilities. • HIV testing (2000): introduction of voluntary counselling and testing centers.	Rolled out later and did not affect specific birth cohorts.
Abortion policy change	• The Penal Code Amendment Act of 11 October 1991 (1991): abortion was de-criminalized within the first 16 weeks of pregnancy under certain circumstances.	Implemented earlier and did not affect specific birth cohorts.
Family planning changes	• National Population Policy (1997): improve quality of life and standard of living of all people; through reduced population growth rate, low fertility, low morbidity and mortality, and a balanced population distribution.	Proposed almost two years after the 1996 reform (in August 1997); unlikely to have affected specific birth cohorts.
Socio-economic reforms	• National Development Plan Number (1991-1997): outline short to medium-term development initiatives for Botswana	Implemented gradually over many years.
Drought relief program	• Botswana's drought relief program (1982-1990): address the loss of livestock and malnutrition, particularly among children sufficient food.	Implemented gradually over many years.

Appendix B

Appendix to Chapter 2

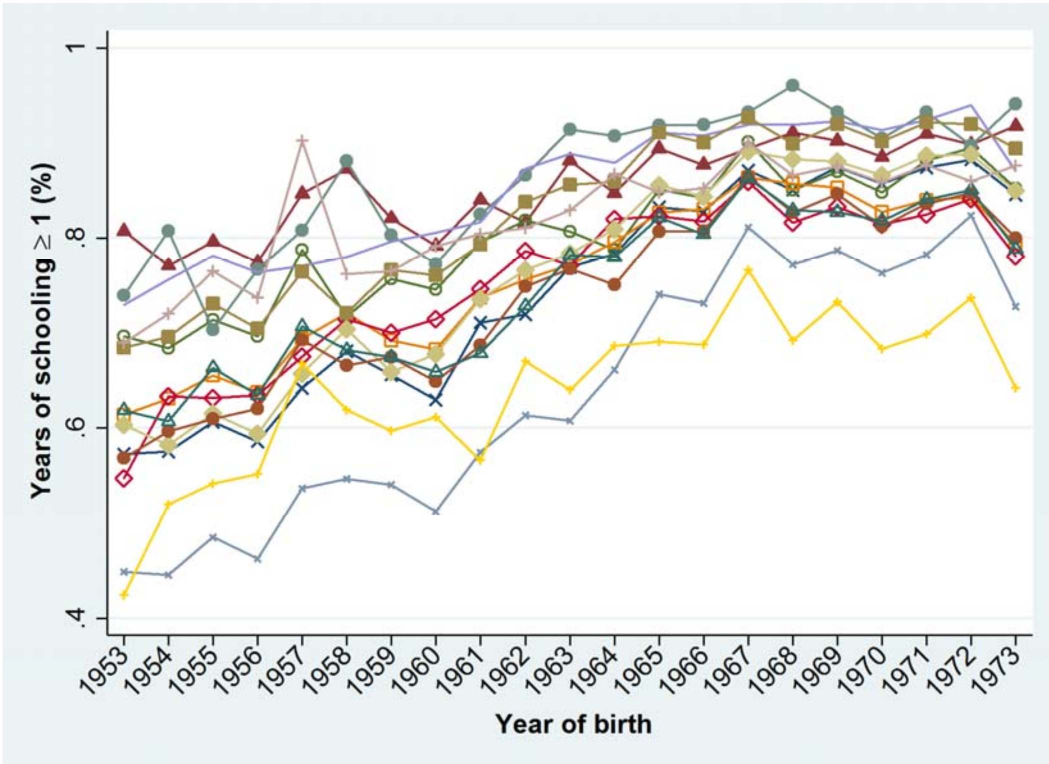
B.1 Supplementary Figures and Tables

Fig. B1: *Educational Attainment by Birth Cohort in Tanzania*



Notes: Figure displays educational attainment by birth cohort in Tanzania with and without education reform (predicted educational attainment with reform: blue diamond markers; predicted educational attainment without reform: red circle markers). Regressions include indicators for five-year age group, an indicator for heap year, indicator for census year, indicators for province of birth, indicator for sex, and the interactions of each covariate with sex. Children born in 1959 and earlier were age 15 or older when the reform was launched, and thus no longer exposed. Children born in 1967 were age 7 when the reform was launched, and thus fully exposed. Children from the 1960 - 1966 cohorts were partially exposed to the reform. Sample includes survey respondents who were Tanzanian citizens and born in Tanzania. Source: Tanzania Census 1988 and 2002.

Fig. B2: *Percentage of Birth Cohort Enrolled in School, by Province of Birth*

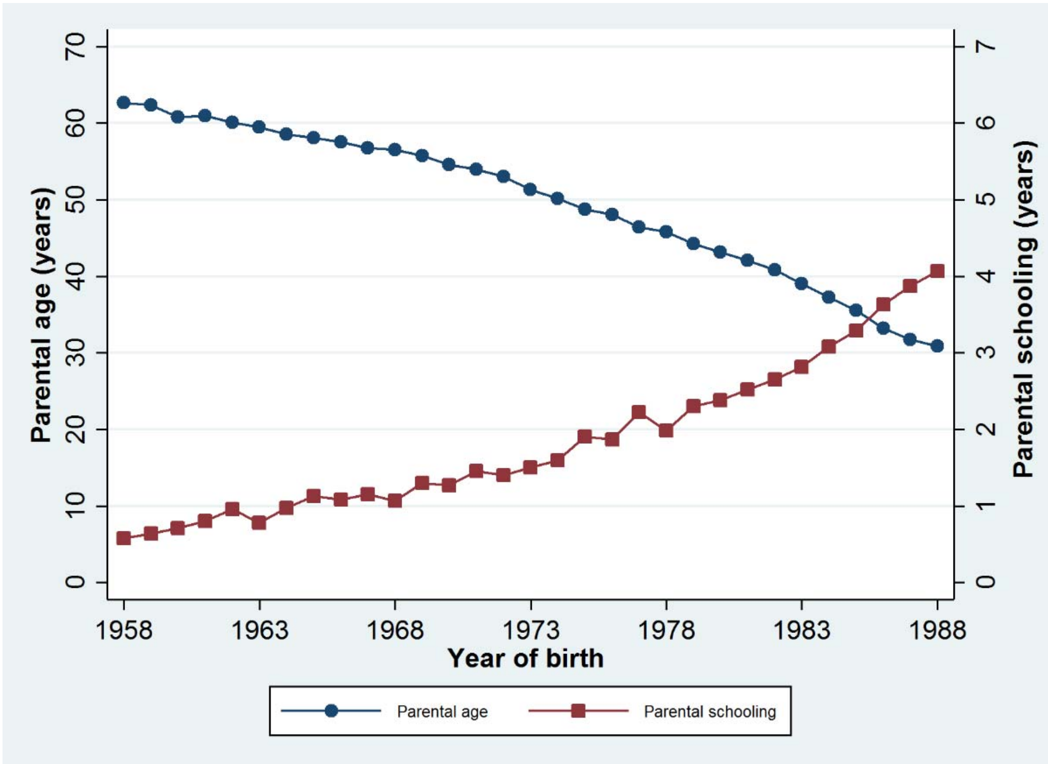


Notes: Figure shows the percentage of each cohort with at least one year of schooling by pairs of neighboring birth provinces using data from the 1988 and 2002 census.

Fig. B3: *Parental Survival, by Children's Years of Schooling*

Notes: Figure shows parental mortality status by a child's educational attainment, separately for mothers and fathers. Educational attainment was defined as total years of schooling completed by the time of the census. Strong protective effects are seen for primary school, with protective effects levelling off after lower secondary school. Sample includes respondents who are Tanzanian citizens, born in Tanzania, and ages at least 20 at the time of the census. Error bars represent 95 % confidence intervals. Source: Tanzania Census 2002.

Fig. B4: *Parental Age and Educational Attainment, by Children's Birth Cohort*



Notes: Figure shows parental age (blue circles) and educational attainment (red squares) for parents co-residing with their children, by year of birth of their youngest child, using data from the 1988 census. Parental educational attainment was defined as total years of schooling completed by the time of the census. Parental age and educational attainment were smooth across birth cohorts of children. Results were similar using alternative specifications.

Table B1: *First Stage Results: Changing Exposure Years*

Dependent Variable	**Schooling (Years)**			
Age Specification	**Age**	**Age, Age^2**	**Age, Age^2, Age^3, Age^4**	**Five-Year Age Group Indicator**
Predictor				
A. Coefficient on Partial Reform Exposure Indicator, YOB 1962 - 1966	1.520***	1.301***	1.298***	1.378***
	(0.007)	(0.008)	(0.008)	(0.008)
Coefficient on Full Reform Exposure Indicator, YOB $\geq$ 1967	1.627***	1.486***	1.491***	1.519***
	(0.009)	(0.009)	(0.009)	(0.009)
B. Coefficient on Partial Reform Exposure Indicator, YOB 1960 - 1964	1.490***	1.244***	1.253***	1.406***
	(0.008)	(0.008)	(0.008)	(0.009)
Coefficient on Full Reform Exposure Indicator, YOB $\geq$ 1965	2.099***	1.874***	1.897***	1.926***
	(0.009)	(0.009)	(0.009)	(0.009)
Observations	2,706,766	2,706,766	2,706,766	2,706,766

OLS models. In addition to age, all regressions include an indicator for heap year, indicator for census year, indicators for province of birth, indicator for sex, and the interactions of each covariate with sex. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights were used. Sample includes Tanzanian citizens born in Tanzania after 1945 and ages at least 15 years old. Source: Tanzania Census 1988 and 2002.

Table B2: *First Stage Results: Alternative Sample Period*

Dependent Variable	**Schooling (Years)**			
Age Specification	**Age**	**Age, Age^2**	**Age, Age^2, Age^3, Age^4**	**Five-Year Age Group Indicator**
Predictor				
A. Sample YOB 1890 - 1987				
Coefficient on Partial Reform Exposure Indicator	1.429***	1.569***	1.199***	1.326***
	(0.007)	(0.007)	(0.007)	(0.008)
Coefficient on Full Reform Exposure Indicator	1.134***	1.610***	1.284***	1.360***
	(0.007)	(0.008)	(0.009)	(0.009)
Observations	3,251,376	3,251,376	3,251,376	3,251,376
B. Sample YOB 1952 - 1987				
Coefficient on Partial Reform Exposure Indicator	1.377***	1.187***	1.129***	1.217***
	(0.008)	(0.008)	(0.008)	(0.009)
Coefficient on Full Reform Exposure Indicator	1.982***	1.746***	1.779***	1.790***
	(0.010)	(0.010)	(0.010)	(0.010)
Observations	2,489,441	2,489,441	2,489,441	2,489,441

OLS models. In addition to age, all regressions include an indicator for heap year, indicator for census year, indicators for province of birth, indicator for sex, and the interactions of each covariate with sex. Full exposure to the reform was defined as 1 if year of birth >= 1967, zero otherwise. Partial exposure to the reform was defined as 1 if year of birth 1960 - 1966, zero otherwise. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights were used. Sample includes Tanzanian citizens born in Tanzania and ages at least 15 years old. Source: Tanzania Census 1988 and 2002.

Table B3: *ITT results: Effect of the Policy Reform on Parental Survival*

Dependent Variable	**Mother Alive (1=yes, 0=no)**			**Father Alive (1=yes, 0=no)**		
Subsample	**Female**	**Male**	**Both Sexes**	**Female**	**Male**	**Both Sexes**
Predictor						
Coefficient on Full Reform Exposure Indicator	0.051***	0.043***	0.047***	0.018***	0.040***	0.029***
	(0.002)	(0.002)	(0.001)	(0.006)	(0.006)	(0.004)
Probability Dependent Variable, Pre-Reform	0.696	0.725	0.710	0.381	0.406	0.393
Observations	1,442,732	1,264,034	2,706,766	958,122	846,076	1,804,198

OLS models controlling for an indicator for five-year age group, indicator for census year (in regressions for Mother Alive), indicator for heap year, indicators for province of birth, and an indicator for partial exposure to the reform (year of birth 1960 - 1966). Regressions for the subsample with both sexes additionally control for an indicator for sex and interactions of each covariate with sex. Full reform exposure was defined as year of birth >= 1967. Father Alive was not available in the 1988 census. Sample includes survey respondents who were citizens of Tanzania, born in Tanzania after 1945, and ages at least 15 years old. No weights were used. Source: Tanzania Census 1988 and 2002.

Table B4: *Parental Survival by Children's Schooling, Post-Reform*

Schooling (Years)	Father Alive (%)	Mother Alive (%)
0	54.9	73.6
1	52.2	73.2
2	55.8	74.6
3	55.9	75.5
4	53.7	74.2
5	54.3	73.3
6	55.7	77.7
7	61.3	80.1
8	61.1	81.2
9	61.7	81.1
10	61.4	83.1
11	66.6	83.6
12	50.1	70.5
13	69.4	83.1
14	65.6	85.8

Schooling was total years of schooling completed by the time of the census. Mother Alive was "mortality status of mother" and Father Alive was "mortality status of father", indicating whether a respondent's biological mother and father were still living at the time of the census. Sample includes survey respondents who were Tanzanian citizens and born in Tanzania between 1967 and 1972 (post-reform). Sample weights used as provided. Source: Tanzania Census 2002.

Table B5: *Logistic Regression Results: Schooling and Parental Survival*

Dependent Variable	**Mother Alive**			**Father Alive**	
Subsample	**Female**	**Male**		**Female**	**Male**
Predictor					
Years of Schooling	1.068***	1.052***		1.054***	1.044***
	(1.066 - 1.069)	(1.050 - 1.053)		(1.053 - 1.056)	(1.043 - 1.046)
Mother's Age: 70-74	0.972***	0.974***	Father's Age: 70-74	0.966***	0.969***
	(0.971 - 0.974)	(0.973 - 0.976)		(0.963 - 0.968)	(0.967 - 0.971)
Mother's Age: 75-79	0.980***	0.982***	Father's Age: 75-79	0.963***	0.960***
	(0.978 - 0.983)	(0.979 - 0.984)		(0.960 - 0.965)	(0.957 - 0.963)
Mother's Age: 80-84	0.980***	0.980***	Father's Age: 80-84	0.990***	0.988***
	(0.978 - 0.982)	(0.978 - 0.981)		(0.988 - 0.991)	(0.987 - 0.990)
Mother's Age: 85-90	0.976***	0.974***	Father's Age: 85-90	0.977***	0.978***
	(0.973 - 0.978)	(0.972 - 0.977)		(0.974 - 0.980)	(0.975 - 0.981)
Mother's Age: 90+	0.972***	0.969***	Father's Age: 90+	0.971***	0.966***
	(0.971 - 0.973)	(0.968 - 0.970)		(0.969 - 0.972)	(0.965 - 0.968)
Mother's Education: Primary	1.026***	1.033***	Father's Education: Primary	1.016***	1.016***
	(1.023 - 1.029)	(1.030 - 1.037)		(1.014 - 1.017)	(1.015 - 1.018)
Mother's Education: Secondary Or More	1.367***	1.261***	Father's Education: Secondary Or More	1.158***	1.149***
	(1.304 - 1.434)	(1.199 - 1.326)		(1.146 - 1.170)	(1.137 - 1.161)
Observations	916,663	795,055		916,547	794,925

All models are logistic regression models and odds ratios are reported with 95% confidence intervals in parentheses. All regressions include indicators for province of birth. A parent's age and education is the proportion of parents in that age or education category in a respondent's birth cohort in the Tanzania Census of 1988. The reference groups for parental age and education in the 1988 census are Age Less Than 70 and Parent's Education: None. *** p<0.01, ** p<0.05, * p<0.1. No weights were used. Sample includes survey respondents who were Tanzanian citizens and born in Tanzania. Source: Tanzania Census 2002.

Table B6: *OLS Results: Survival Rates*

Dependent Variable: 14-Year Survival Rate	**Female**	**Male**
Predictor		
Coefficient on Full Reform Exposure	0.136***	0.134***
	(0.001)	(0.001)
Control: Cohort Education, Census 1988	✓	✓
Sample: Census 2002	✓	✓
Observations	656,065	590,274
Coefficient on Full Reform Exposure	0.046***	0.036***
	(0.000)	(0.000)
Control: Cohort Education, Census 1988	✓	✓
Sample: Both Census Years	✓	✓
Observations	1,301,971	1,165,580

OLS models. Survival rates were calculated as the number of people alive in 2002 divided by the number of people alive in 1988 for each birth cohort. Cohort Reform Exposure was calculated as the percentage of the cohort having a child post reform based on an indicator for full parental exposure to the reform (1 if year of birth of eldest child was >= 1967, zero otherwise). All models control for indicators for five-year age group, indicator for heap year, indicators for province of birth, and an indicator for parental partial exposure to the reform (year of birth of eldest child 1960 - 1966). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights were used. Sample includes survey respondents who were citizens of Tanzania and born in Tanzania. Source: Tanzania Census 1988 and 2002.

Table B7: *2SLS Results: Controlling for Labor Force Participation*

Dependent Variable	**Mother Alive (1=yes, 0=no)**			**Father Alive (1=yes, 0=no)**		
Subsample	**Female**	**Male**	**Both Sexes**	**Female**	**Male**	**Both Sexes**
Coefficient on Schooling (Years)	0.025***	0.039***	0.031***	0.022**	0.090***	0.045***
	(0.001)	(0.001)	(0.001)	(0.011)	(0.021)	(0.010)
Additional Control						
Labor Force Participation (1=yes, 0=no)	✓	✓	✓	✓	✓	✓
Observations	1,441,335	1,261,913	2,703,248	956,975	844,262	1,801,237

Regression results from 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. The instruments were an indicator for partial exposure to the reform (YOB 1960 - 1966) and an indicator for full exposure to the reform (YOB >= 1967). All models control for indicators for five-year age group, indicator for census year (in regressions for Mother Alive), indicator for heap year, and indicators for province of birth. Regressions for the subsample with both sexes additionally control for an indicator for sex and interactions of each covariate with sex. F-statistics were 16,902 (both census years) and 95 (census 2002) for the female subsample; 6,204 (both census years) and 39 (census 2002) for the male subsample; and 21,010 (both census years) and 131 (census 2002) in the pooled sample. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents who were citizens of Tanzania, born in Tanzania after 1945, and ages at least 15 years old. No weights were used. Source: Tanzania Census 1988 and 2002.

Table B8.A: *2SLS Results: Labor Force Participation, Literacy, Marital Status, Spousal Schooling, and Co-Residence with Parents*

Dependent Variable	**Labor Force Participation (1=yes, 0=no)**	**Literacy (1=yes, 0=no)**	**Married (1=yes, 0=no)**	**Spousal Schooling (Years)**	**Mother Present in Respondent's Household (1=yes, 0=no)**	**Father Present in Respondent's Household (1=yes, 0=no)**
Coefficient on Schooling (Years)						
Both Sexes	0.020***	0.099***	-0.020***	0.710***	0.004***	0.012***
	(0.001)	(0.000)	(0.001)	(0.007)	(0.001)	(0.000)
Observations	2,703,248	2,706,714	2,706,763	1,163,452	2,706,765	2,706,766
F-Statistic	20,537	20,546	20,548	8,131	20,547	20,548
Female	0.015***	0.110***	-0.004***	0.589***	0.002***	0.007***
	(0.001)	(0.000)	(0.001)	(0.007)	(0.001)	(0.001)
Observations	1,441,335	1,442,704	1,442,731	644,241	1,442,731	1,442,732
F-Statistic	16,678	16,678	16,680	7,325	16,680	16,680
Male	0.021***	0.077***	-0.039***	1.067***	0.008***	0.018***
	(0.001)	(0.001)	(0.001)	(0.015)	(0.001)	(0.001)
Observations	1,261,913	1,264,010	1,264,032	519,211	1,264,034	1,264,034
F-Statistic	6,054	6,058	6,059	2,910	6,058	6,058

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. The instruments were an indicator for partial exposure to the reform (YOB 1960 - 1966) and an indicator for full exposure to the reform (YOB >= 1967). All models included the following controls: indicators for five-year age group, indicator for census year, indicator for heap year, indicators for province of birth, indicators for current region, and the interaction of current region and census year. Regressions for the subsample with both sexes additionally control for sex and the interactions of each covariate with sex. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights were used. Sample includes survey respondents who were citizens of Tanzania, born in Tanzania after 1945, and ages at least 15 years old. No weights were used. Source: Tanzania Census 1988 and 2002.

Table B8.B: *2SLS Results: Number of Co-Resident Family Members, and Household Access to Basic Utilities*

Dependent Variable	Number of Own Family Members in Houshold	Toilet or Latrine, Conditional on Parental Household Head (1=yes, 0=no)	Running Water, Conditional on Parental Household Head (1=yes, 0=no)	Electricity, Conditional on Parental Household Head (1=yes, 0=no)
Coefficient on Schooling (Years)				
Both Sexes	0.118***	0.001	0.008***	0.003**
	(0.006)	(0.001)	(0.002)	(0.001)
Observations	2,706,766	575,216	575,440	575,437
F-Statistic	20,548	2,080	2,080	2,080
Female	0.026***	0.000	0.011**	0.003
	(0.006)	(0.003)	(0.004)	(0.002)
Observations	1,442,732	308,574	308,683	308,682
F-Statistic	16,680	1,938	1,939	1,940
Male	0.221***	0.003**	0.005**	0.003**
	(0.010)	(0.001)	(0.002)	(0.001)
Observations	1,264,034	266,642	266,757	266,755
F-Statistic	6,058	436	436	436

All regressions are 2SLS models, in which exposure to the reform was used as an instrument for years of schooling. The instruments were an indicator for partial exposure to the reform (YOB 1960 - 1966) and an indicator for full exposure to the reform (YOB >= 1967). All models included the following controls: indicators for five-year age group, indicator for census year, indicator for heap year, indicators for province of birth, indicators for current region, and the interaction of current region and census year. Regressions for the subsample with both sexes additionally control for sex and the interactions of each covariate with sex. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No weights were used. Sample includes survey respondents who were citizens of Tanzania, born in Tanzania after 1945, and ages at least 15 years old. No weights were used. Source: Tanzania Census 1988 and 2002.

Appendix C

Appendix to Chapter 3

C.1 Childhood Nutrition Context

Stunting, in particular, is highly endemic in Zimbabwe. The prevalence of stunting among children under 5 is non-monotonic, with prevalence peaking at 49% for children in the age range of 24 – 35 months, before declining somewhat as children approach their 5th birthday (26% of children age 48 – 59 months are stunted). The prevalence of wasting in Zimbabwe, on the other hand, is highest among children age 9 – 11 months (8%). The prevalence of underweight is highest among children age 18 – 23 months (16%) (Demographic and Health Surveys 2011). Figure C4 displays childhood undernutrition by age in my study sample. Boys were slightly more likely to suffer from undernutrition than girls across all three indices. The prevalence of stunting and underweight in Zimbabwe has increased between the 1999 and 2005-06 DHS surveys and decreased between the 2005 – 06 and 2010 – 11 DHS surveys; in contrast, the prevalence of wasting has declined since the 1999 DHS survey. Additionally, overweight has become a major concern in Zimbabwe with 6% of children under 5 overweight, and 31% of women and 9% of men ages 15 – 49 either overweight ($BMI \geq 25$ m/kg^2) or obese ($BMI \geq 30$ m/kg^2) (Garrett & Ruel 2005). The co-existence of malnutrition among children and their parents has important nutrition policy implications: households at risk of both undernutrition and overweight may require more complex prevention interventions.

C.2 Education System and Reform

Zimbabwe's public education system is based on the "7-4-2-3" system, including seven years of primary school, four years of lower secondary school ("Ordinary Level"; General Certificate of Education), two years of higher secondary school ("Advanced Level"; General Certificate of Education Advanced Level), and three years of tertiary education (UNESCO 2010). In 1980, following independence, Zimbabwe rapidly expanded access to secondary schools for black Zimbabweans. In particular, the government initiated a large secondary school construction program, focusing on rural areas. While expansions in education occurred at both the primary and secondary levels, the expansion of secondary schools was by far the most dramatic (Lemon 1995, Ansell 2002). As a result of the reform, enrollment in secondary school more than doubled, increasing from 74,746 in 1980 to 149,018 in 1981 (UNESCO 2015). Previous studies have found that the reform led to a large increase in secondary school attainment (Agüero & Bharadwaj 2014, Grépin & Bharadwaj 2015). This 'natural experiment' provides an opportunity to estimate the causal impact of parental schooling on undernutrition in their children, by comparing children of parental birth cohorts exposed to the reform versus those unexposed. While independence might have also led to other policy reforms, the fundamental assumption underlying this study is that these other changes did not affect parental cohorts differentially around the year of birth cutoffs.

C.3 References for Appendix C

Ansell, N (2002). Secondary Education Reform in Lesotho and Zimbabwe and the Needs of Rural Girls: Pronouncements, policy and practice. *Comparative Education*. 38(1):91-112.

Agüero, J.M., Bharadwaj, P (2014). Do the More Educated Know More about Health? Evidence from Schooling and HIV Knowledge in Zimbabwe. *Economic Development and Cultural Change*. 62(3):489-517.

Demographic and Health Surveys final reports Zimbabwe (2011). Demographic and Health Surveys.

Garrett, J.L., Ruel, M.T. (2005). Stunted child-overweight mother pairs: prevalence and association with economic development and urbanization. *Food and Nutrition Bulletin*. 26(2):209-21.

Grépin, K.A., Bharadwaj, P. (2015) Maternal education and child mortality in Zimbabwe. *Journal of Health Economics*. 44:97-117.

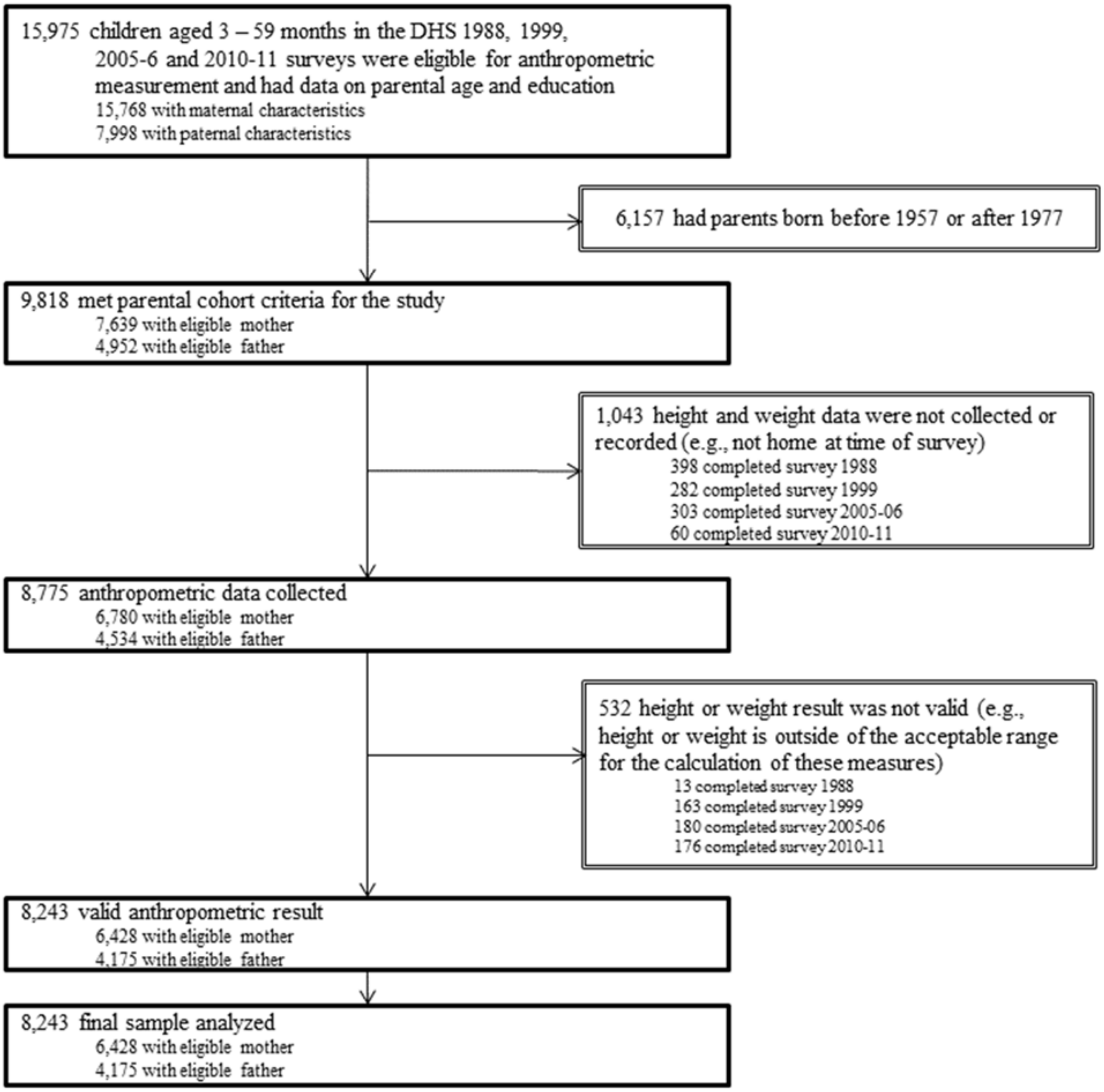
Lemon, A. (1995) Education in Post-apartheid South Africa: Some lessons from Zimbabwe. *Comparative Education*. 1995;31(1):101-14.

UNESCO Institute for Statistics (2015), Zimbabwe Education Statistics.

UNESCO (2010). World Data on Education VII Ed. 2010/11

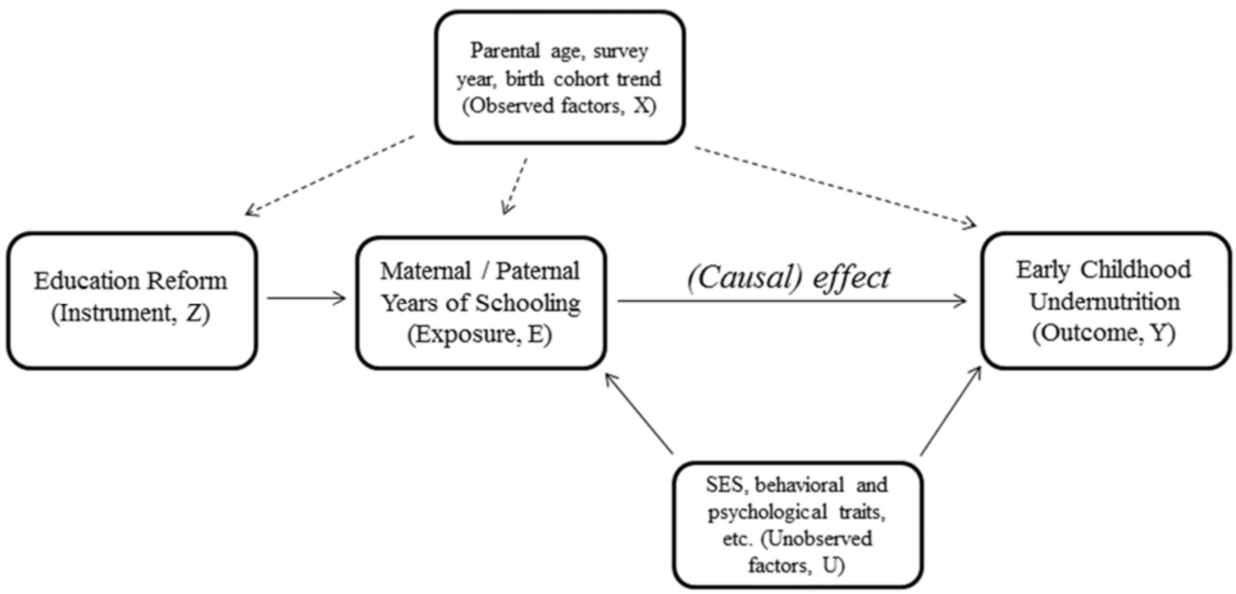
C.4 Supplementary Figures and Tables

Fig. C1: *Study Participants*



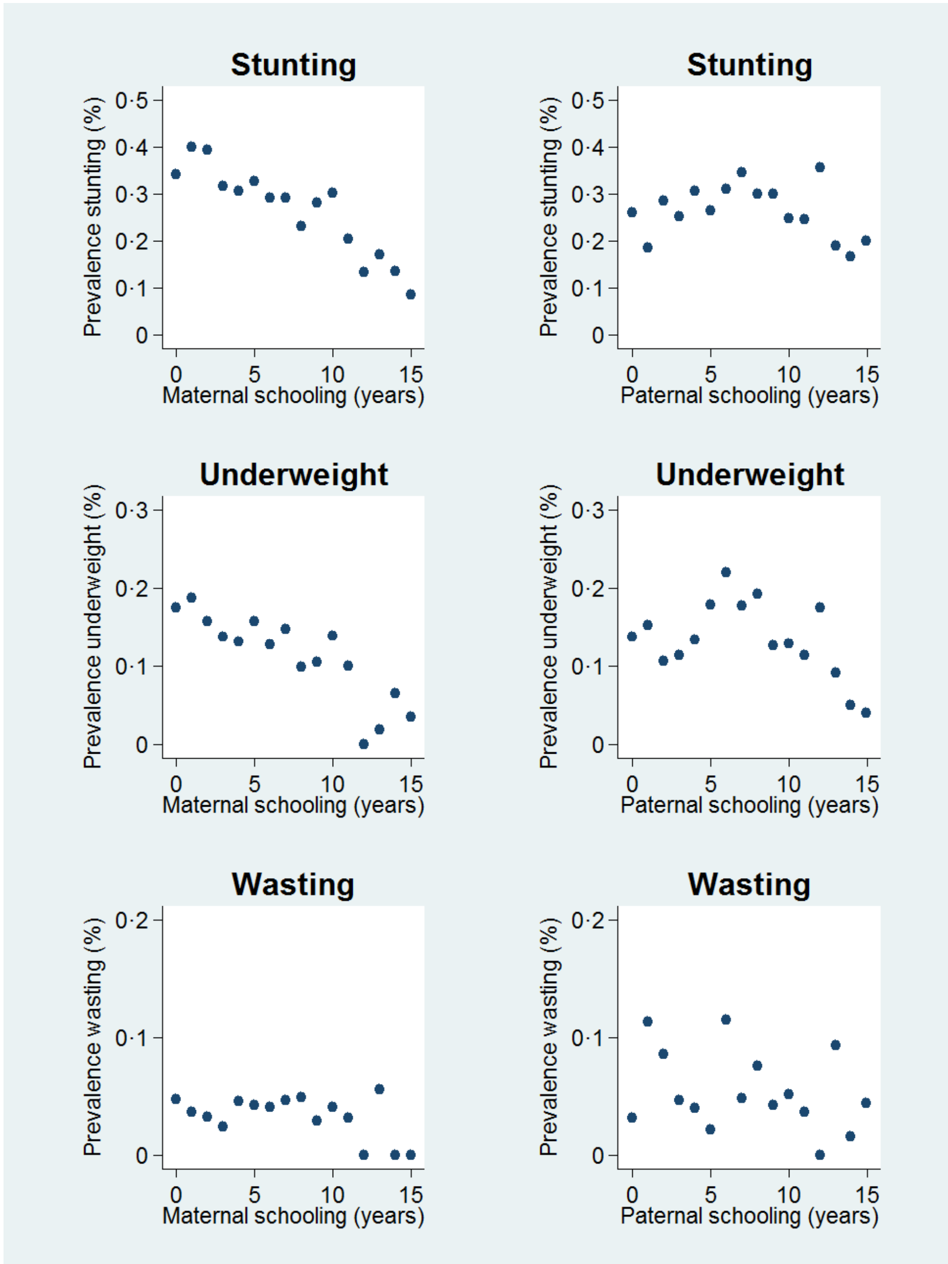
Notes: Diagram showing the flow of study participants through the Zimbabwe Demographic and Health Surveys of 1988, 1999, 2005-6 and 2010-11.

Fig. C2: *Assumptions Underpinning the Study*



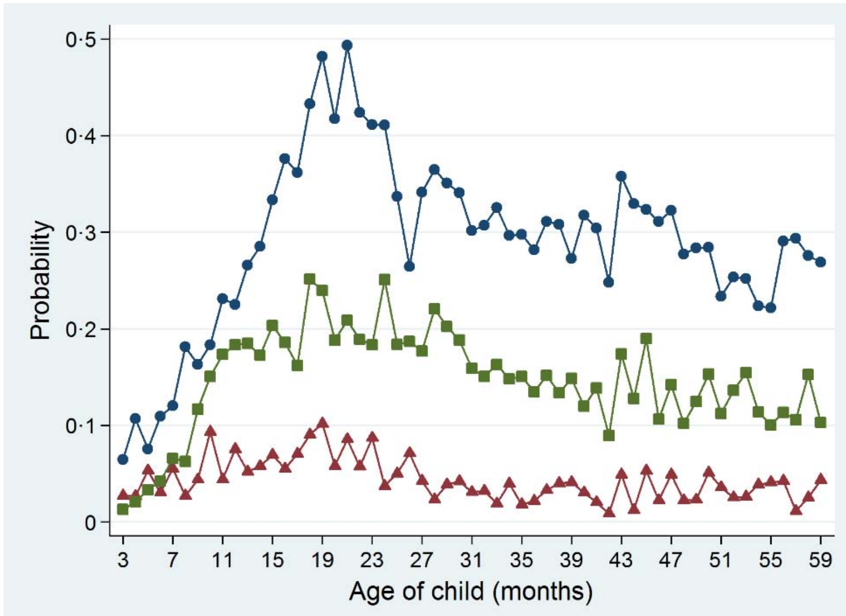
Notes: Directed acyclic graph illustrating the instrumental variable assumptions for causal interpretation in my study. Conditional on X, Z is a valid instrument if Z causally affects E, Z is uncorrelated with U, and Z affects Y only through E. Under the assumption that Z only affects E in one direction, two-stage least squares identifies a local average treatment effect.

Fig. C3: *Undernutrition in Children by Parental Years of Schooling in Zimbabwe*



Notes: Figure shows undernutrition in children by parental years of schooling completed, separately for maternal and paternal schooling. Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 – 1977. Survey weights used as provided. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

Fig. C4: *Childhood Undernutrition by Age in Zimbabwe*



Notes: Figure shows early childhood undernutrition by age in months, defined as stunting (blue circles), underweight (green squares), and wasting (red triangles). Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 – 1977. Survey weights used as provided. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

Table C1: *OLS First Stage Regression Results: Effect of the Policy Reform on Parental Educational Attainment*

Outcome	**Years of schooling**	**1+ Years**	**7+ Years**	**9+ Years**	**11+ Years**	**13+ Years**	**14+ Years**	**9+ Years (Placebo)**
Subsample: mothers								
Effect estimate: reform indicator (SE)	2.497***	0.135***	0.264***	0.295***	0.174***	0.019*	0.019**	0.017
	(0.181)	(0.015)	(0.023)	(0.025)	(0.020)	(0.010)	(0.009)	(0.026)
Mean outcome	7.0	0.900	0.662	0.363	0.180	0.033	0.028	-
No. of mothers	12,925	12,925	12,925	12,925	12,925	12,925	12,925	13,674
R-squared	0.155	0.071	0.148	0.107	0.104	0.011	0.011	0.150
Subsample: fathers								
Effect estimate: reform indicator (SE)	2.393***	0.037***	0.176***	0.389***	0.306***	0.080***	0.073***	-0.053
	(0.233)	(0.012)	(0.024)	(0.033)	(0.031)	(0.022)	(0.020)	(0.035)
Mean outcome	9.0	0.970	0.845	0.593	0.345	0.096	0.079	-
No. of fathers	7,134	7,134	7,134	7,134	7,134	7,134	7,134	6,914
R-squared	0.098	0.036	0.087	0.095	0.186	0.015	0.014	0.033

Regression results from OLS models controlling for an indicator for parental age, a linear term for parental year of birth (YOB), an indicator for survey year, and an indicator for partial exposure to the reform (parental YOB 1963 - 1966). Full exposure to the reform was defined as 1 if parental YOB >= 1967, zero otherwise. Robust standard errors (SE) in parentheses. In model 8, I shifted the reform cutoff for treatment and control cohorts by 5 years forward, using the sample of respondents born between 1962 and 1982. *** p<0.01, ** p<0.05, * p<0.1. Sample includes survey respondents born between 1957 and 1977 who were parents and lived with their child in the same household. No weights were used. Source: Zimbabwe DHS 1988, 1994, 1999, 2005-6 and 2010-11.

Table C2: *OLS Intention-to-Treat Regression Results*

Exposure	**Maternal**			**Paternal**		
Outcome	**Stunted**	**Underweight**	**Wasted**	**Stunted**	**Underweight**	**Wasted**
Effect estimate: reform indicator (SE)	0.041 (0.039)	0.036 (0.030)	-0.018 (0.018)	0.055 (0.042)	0.002 (0.033)	-0.016 (0.020)
No. of children	6,428	6,428	6,428	4,175	4,175	4,175
R-squared	0.009	0.010	0.014	0.009	0.009	0.013

OLS regressions results controlling for an indicator for parental age, a linear trend in parental year of birth (YOB), an indicator for survey year, and an indicator for partial exposure to the reform (YOB 1963 - 1966). Full exposure to the reform was defined as 1 if parental YOB >= 1967, zero otherwise. Robust standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. No weights were used. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

Table C3: *Sensitivity Analyses Regression Results: Region Indicators, Quadratic in Parental Year of Birth, Sample Weights, Probit, Controlling for Child Age, Sex, and Birth Order, and Alternative Sample Specifications and Identification Strategy*

Model	(1) ITT	(2) ITT	(3) ITT	(4) ITT	(5) ITT	(6) ITT	(7) ITT	(8) ITT	(9) ITT	(10) ITT	(11) 2SLS (IV)
Sensitivity analysis	Region	Quadratic year of birth	Sample weights	Probit	Child age, sex, birth order	3-year cohort window	5-year cohort window	Excluding partially exposed	Parental schooling 7+ years	Full sample	Diff-in-diff
Exposure	Reform indicator (SE)	Reform indicator (SE)	Reform indicator (SE)	Reform indicator (SE)	Reform indicator (SE)	Reform indicator (SE)	Reform indicator (SE)	Reform indicator (SE)	Reform indicator (SE)	Reform indicator (SE)	Years of schooling (SE)
						Maternal exposure					
Outcome											
Stunted	0.040	0.064	0.033	0.043	0.051	0.072	0.144	0.045	0.035	0.012	0.012
	(0.038)	(0.058)	(0.043)	(0.038)	(0.039)	(0.107)	(0.123)	(0.041)	(0.049)	(0.021)	(0.029)
Underweight	0.036	0.032	0.022	0.034	0.047	-0.010	-0.015	0.033	0.069*	0.015	0.009
	(0.030)	(0.043)	(0.033)	(0.029)	(0.031)	(0.066)	(0.082)	(0.032)	(0.036)	(0.016)	(0.022)
Wasted	-0.020	-0.015	-0.029	-0.022	-0.007	-0.004	0.001	-0.035*	-0.003	-0.016*	0.017
	(0.018)	(0.027)	(0.020)	(0.017)	(0.019)	(0.040)	(0.051)	(0.020)	(0.021)	(0.009)	(0.013)
No. of children	6,428	6,428	6,428	6,428	6,182	1,598	2,799	5,469	4,220	13,461	6,428
						Paternal exposure					
Outcome											
Stunted	0.056	0.074	0.070	0.050	0.062	0.172	0.158	0.046	0.039	0.019	-0.016
	(0.042)	(0.069)	(0.046)	(0.041)	(0.042)	(0.115)	(0.128)	(0.042)	(0.034)	(0.031)	(0.035)
Underweight	0.005	-0.001	-0.006	-0.002	0.009	-0.025	-0.054	-0.009	-0.012	0.012	-0.019
	(0.032)	(0.052)	(0.035)	(0.032)	(0.033)	(0.076)	(0.089)	(0.033)	(0.026)	(0.024)	(0.027)
Wasted	-0.018	-0.007	-0.018	-0.025	-0.019	-0.046	-0.060	-0.022	-0.006	0.000	-0.020
	(0.020)	(0.031)	(0.020)	(0.021)	(0.020)	(0.029)	(0.039)	(0.020)	(0.016)	(0.015)	(0.020)
No. of children	4,175	4,175	4,175	4,175	3,879	1,083	1,848	3,592	3,580	6,832	4,175

Models 1 - 10 are regressions results from intention-to-treat (ITT) OLS models, controlling for an indicator for parental age, a linear trend in parental year of birth (YOB), an indicator for survey year, and an indicator for partial exposure to the reform (parental YOB 1963 - 1966). Full exposure to the reform was defined as 1 if parental YOB >= 1967, zero otherwise. Model 2 additionally controls for parental YOB squared. Model 4 is a Probit model. Model 5 additional controls for fixed effects in child age, sex, and birth order. Model 11 is a two-stage least squares (2SLS) model, where the instrumental variable (IV) was the interaction beween a reform indicator and a continous variable for the proportion of respondents with no schooling by region in the pre-reform period, while controlling for the main effects of each variable. Robust standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No sample weights were used, except in model 3. Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

Table C4: *2SLS Regression Results Accounting for Measurement Error in Parental Age*

Exposure	Maternal schooling			Paternal schooling		
Outcome	**Stunted**	**Underweight**	**Wasted**	**Stunted**	**Underweight**	**Wasted**
Additional control for parental heap year indicator (A)						
Effect estimate:	0.015	0.016	-0.005	0.017	0.005	-0.007
years of schooling (SE)	(0.014)	(0.011)	(0.007)	(0.017)	(0.013)	(0.008)
F-statistic	49.6	49.6	49.6	33.3	33.3	33.3
No. of children	6,424	6,424	6,424	4,150	4,150	4,150
Excluding parental ages that are multiples of five (B)						
Effect estimate:	0.011	0.011	-0.004	0.017	0.006	-0.008
years of schooling (SE)	(0.015)	(0.012)	(0.007)	(0.017)	(0.013)	(0.008)
F-statistic	39.5	39.5	39.5	30.2	30.2	30.2
No. of children	5,040	5,040	5,040	3,248	3,248	3,248

Regression results from a 2SLS model, in which exposure to the reform was used as an instrument for parental years of schooling. All models control for an indicator for parental age, a linear trend in parental YOB, and an indicator for survey year. Panel A additionally controls for an indicator for parental heap year. Panel B excludes respondents with parental ages that are multiples of five. Since the surveys were conducted in 6 different years, I still have all birth cohorts. The instruments were an indicator for partial exposure to the reform (parental YOB 1963 - 1966) and an indicator for full exposure to the reform (parental YOB >= 1967). Robust standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. No weights were used. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

Table C5: *OLS and 2SLS Regression Results Using New WHO Standards*

Exposure	**Maternal schooling**			**Paternal schooling**		
Outcome	**Stunted, WHO 2006#**	**Underweight, WHO 2006#**	**Wasted, WHO 2006#**	**Stunted, WHO 2006#**	**Underweight, WHO 2006#**	**Wasted, WHO 2006#**
Model 1: OLS						
Effect estimate: years of schooling (SE)	-0.016*** (0.002)	-0.009*** (0.001)	-0.003*** (0.001)	-0.011*** (0.002)	-0.007*** (0.002)	-0.003** (0.001)
Model 2: 2SLS (IV)						
Effect estimate: years of schooling (SE)	0.035** (0.015)	0.018* (0.010)	-0.008 (0.007)	0.001 (0.017)	0.007 (0.012)	0.003 (0.008)
F-statistic	49.8	49.8	49.8	34.6	34.6	34.6
Mean outcome	0.326	0.104	0.047	0.325	0.110	0.055
No. of children	6,428	6,428	6,428	4,175	4,175	4,175

Model 1 is an ordinary least squares (OLS) linear probability model. Model 2 is a two-stage least squares (2SLS) linear probability model, in which exposure to the reform was used as an instrumental variable (IV) for parental years of schooling. All models included the following controls: an indicator for parental age, a linear trend in parental year of birth (YOB), and an indicator for survey year. The instruments were an indicator for partial exposure to the reform (YOB 1963 - 1966) and an indicator for full exposure to the reform (YOB >= 1967). Robust standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. #Undernutrition criteria as newly defined by the WHO in 2006 (Borghi et al., 2006). Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. No weights were used. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

Table C6: *OLS and 2SLS Regression Results: Effect of Parental Schooling on HAZ, WAZ, and WHZ*

Exposure	**Maternal schooling**			**Paternal schooling**		
Outcome	**Height-for-age z-score**	**Weight-for-age z-score**	**Weight-for-height z-score**	**Height-for-age z-score**	**Weight-for-age z-score**	**Weight-for-height z-score**
Model 1: OLS						
Effect estimate: years of schooling (SE)	0.053***	0.066***	0.045***	0.039***	0.050***	0.036***
	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)
Model 2: 2SLS (IV)						
Effect estimate: years of schooling (SE)	-0.024	0.006	0.037	0.025	0.079*	0.081*
	(0.045)	(0.036)	(0.038)	(0.054)	(0.043)	(0.045)
F-statistic	49.6	49.6	49.6	33.3	33.3	33.3
Mean outcome	-1.25	-0.81	-0.02	-1.17	-0.82	-0.09
No. of children	6,428	6,428	6,428	4,175	4,175	4,175

Model 1 is an ordinary least squares (OLS) linear probability model. Model 2 is a two-stage least squares (2SLS) linear probability model, in which exposure to the reform was used as an instrumental variable (IV) for parental years of schooling. All models included the following controls: an indicator for parental age, a linear trend in parental year of birth (YOB), and an indicator for survey year. The instruments were an indicator for partial exposure to the reform (parental YOB 1963 - 1966) and an indicator for full exposure to the reform (parental YOB >= 1967). Robust standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. No weights were used. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.

Table C7: *OLS and 2SLS Regression Results: Effect of Parental Schooling on Severe Malnutrition and Overweight*

Exposure	**Maternal schooling**				**Paternal schooling**			
Outcome	**Stunted, severe**	**Underweight, severe**	**Wasted, severe**	**Overweight**	**Stunted, severe**	**Underweight, severe**	**Wasted, severe**	**Overweight**
Model 1: OLS								
Effect estimate:	-0.005***	-0.002***	0.000	0.005***	-0.003**	-0.001	0.000	0.003**
years of schooling (SE)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Model 2: 2SLS (IV)								
Effect estimate:	0.009	-0.005	0.002	0.006	0.006	-0.003	0.003	0.010
years of schooling (SE)	(0.009)	(0.005)	(0.003)	(0.007)	(0.010)	(0.005)	(0.004)	(0.008)
F-statistic	49.6	49.6	49.6	49.6	33.3	33.3	33.3	33.3
Mean outcome	0.091	0.021	0.009	0.050	0.089	0.022	0.010	0.047
No. of children	6,428	6,428	6,428	6,428	4,175	4,175	4,175	4,175

Model 1 is an ordinary least squares (OLS) linear probability model. Model 2 is a two-stage least squares (2SLS) linear probability model, in which exposure to the reform was used as an instrumental variable (IV) for parental years of schooling. All models included the following controls: an indicator for parental age, a linear trend in parental year of birth (YOB), and an indicator for survey year. Stunting, underweight, and wasting defined by Z scores of less than –3 were classified as "severe". Overweight was defined by weight-for-height Z scores above 2. The instruments were an indicator for partial exposure to the reform (parental YOB 1963 - 1966) and an indicator for full exposure to the reform (parental YOB >= 1967). Robust standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample includes children ages between 3 and 59 months at the time of the survey, and living with a parent born 1957 - 1977. No weights were used. Source: Zimbabwe DHS 1988, 1999, 2005-6 and 2010-11.