Three Essays on Schooling and Health in Indonesia.
Assessing the Effects of Family Planning on Fertility and of Supply-Side Education Programmes on BMI, Schooling Attainment, and Wages

Gunilla Pettersson

Submitted for the degree of Doctor of Philosophy

Department of Economics

University of Sussex

January 2012

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature: Gunilla Pettersson

UNIVERSITY OF SUSSEX

Gunilla Pettersson, Doctor of Philosophy

Three Essays on Schooling and Health in Indonesia. Assessing the Effects of Family Planning on Fertility and of Supply-Side Education Programmes on BMI, Schooling Attainment, and Wages

SUMMARY

In 1969, Indonesia established a national family planning programme and total fertility has declined rapidly since but there is little consensus over the relative contribution of family planning to the observed decline. The first chapter constructs a new measure of family planning exposure to examine the role of family planning in reducing fertility. The causal effects of infant mortality is also examined based on a new instrumental variable, water supply and sanitation programme exposure, and that of schooling using father's schooling as an instrument. The findings strongly indicate that family planning contributes to lower fertility together with reductions in infant deaths and improvements in women's schooling, and that the effects of family planning and decreases in infant mortality are larger than that of schooling. In 2002, nearly one-in-ten men and more than one-in-five women in Indonesia were overweight and noncommunicable diseases had become the main cause of death but there exists no evidence on the causal effect of schooling on BMI for developing countries. The second chapter assesses whether more schooling causes healthier BMI in Indonesia by using two instrumental variables to capture exogenous variation in schooling. The first instrument takes advantage of the primary school construction programme (SD INPRES) in the 1970s; the second instrument is father's schooling. Two results stand out: more schooling causes higher BMI for men and there is no causal effect of schooling on BMI for women. This chapter also provides some very preliminary evidence that the shift from blue collar to white collar and service sector occupations is one contributing factor to why more schooling increases BMI for men. The third chapter also uses the SD INPRES programme but to examine the effect of increased school supply on schooling attainment: overall, by gender, and by socioeconomic background. It also constructs a new SD INPRES programme exposure variable as an instrument for schooling to assess the causal effect of schooling on wages. The results strongly suggest that the SD INPRES programme increased schooling for men and women but that women benefited more as did individuals from less advantageous socioeconomic backgrounds. More schooling also causes higher wages and there appears to be an added positive effect for women through the additional schooling induced by the SD INPRES programme.

Acknowledgements

I want to thank my supervisor Andrew Newell for providing continuous support, steady guidance, and astute insights during the course of my PhD research. He encouraged my good ideas and tactfully discouraged the not so good ones and cheered me up by talking about completely irrelevant matters as necessary. I also had some excellent conversations with my second supervisor Barry Reilly, although admittedly some of them were about football rather than economics.

Two of my good friends who also happen to be brilliant economists, Erwin Tiongson and Daniel Monchuk, allowed me to bounce ideas of them and answered questions without a moment of hesitance. Sami Bazzi, a real Indonesia expert who I am fortunate to know, assisted in obtaining some hard-to-get data and background documents. Esther Duflo kindly provided some of the administrative data for which I am very grateful. I also thank Hans Werthén Fonden for providing the scholarship that enabled me to do my PhD.

I would not have persisted throughout the PhD process had it not been for my wonderful mother and amazing friends who were always there to cheer me on. So thank you mum, and in no particular order, Maria, Victoria, Nithin, Ellen, Kajsa, Michael, Elizabeth, Göran and Henrik. Finally, thank you to Nancy Birdsall who inspired and encouraged me to do the PhD in the first place.

Table of Contents

List of Tables

List of Figures

List of Annex Tables

List of Annex Figures

Acronyms and Glossary

BKKBN	National Family Planning Coordinating Board
BMI	Body mass index
Bupati	District chief executive
DID	Difference-in-difference estimator
DMCI	Differenced Mean Closed Interval to the Next Birth
FP	Family planning
FPMCH	Outreach Family Planning and Health Services project in Matlab, Bangladesh
GED	General Educational Development High School Equivalency Diploma
GOI	Government of Indonesia
IFLS	Indonesia Family Life Survey
Kabupaten	District
Kecamatan	Sub-district
NFPP	National Family Planning Programme
Posyandu	Community sponsored health post that offers family planning
PPR	Parity Progression Ratio
Puskemas	Health clinic
Repelita	Five-year development plan (Rencana Pembangunan Lima Tahun)
Repelita I	National five-year development plan 1968/69-1973/74
Repelita II	National five-year development plan 1973/74-1978/79
SD INPRES	Presidential Instruction Primary School Programme (Sekolah Dasar)
SUPAS	Intercensal survey of Indonesia
WHO	World Health Organization
WSSP	Water Supply and Sanitation Programme

Chapter 1: Fertility Decline in Indonesia: Family Planning, Infant Mortality, and Schooling Matter.

Abstract

Lower fertility has several benefits at the individual level: fewer births are associated with better health and schooling outcomes for women and their children, with higher household income and wealth, and with improved labour market opportunities for women. Indonesia established a national family planning programme in 1969 and total fertility has declined rapidly since but there is little consensus over the relative contribution of family planning to the observed decline. This chapter empirically examines the effectiveness of family planning in reducing total fertility at the individual level in Indonesia and the findings strongly suggest that family planning contributed to the fertility decline. The causal effect of infant mortality on total fertility is also assessed using a public water supply and sanitation programme implemented in staggered fashion across districts combined with the timing of births to capture exogenous variation in infant mortality. The results strongly indicate that increases in the number of infant deaths cause higher total fertility. This chapter also assesses the causal effect of women's schooling on fertility using the schooling of fathers of the women who give birth as an instrument and finds that more schooling causes lower total fertility. Overall, this chapter shows that family planning contributes to lower fertility together with reductions in infant deaths and improvements in women's schooling, and that the effects of family planning and decreases in infant mortality on fertility are larger than that of women's schooling, which has implications for public policy aimed at reducing fertility.

1.1 Introduction

Numerous developing countries have implemented family planning programmes since the 1960s with the objective of reducing fertility, in part to help promote economic growth and development through a number of channels (Joshi 2011, Schultz 2008). Over the period 1960-2000, the proportion of married women in developing countries using birth control increased from about 10 percent to 60 percent and mean fertility nearly halved from six to three births per woman (Cleland, Bernstein et al. 2006).

Lower fertility has benefits at the individual level. Fewer births are associated with better health outcomes for women and children, for instance, through longer spacing between births, which tends to reduce both maternal and infant mortality (Das Gupta, Boongarts et al. 2011). Lower fertility is also associated with higher household income and wealth (Schultz, Joshi 2007), which allows parents to spend more on the schooling, health, and nutrition of each child with positive implications for their schooling, labour market, and health outcomes (Birdsall 1980, Rosenzweig, Wolpin 1980, Schultz 1997). The ability to control the number of births improves women's opportunities in the labour market and in almost all countries women's labour force participation has risen as fertility has declined (Cleland 2002).

Family planning programmes aim to give women in particular greater control over family formation decisions and the ability to gain some of the benefits of lower fertility. For instance, such programmes can improve women's opportunity to obtain schooling by making it easier to delay the birth of their first child (Joshi 2011).

However, over the last decade and half attention has shifted away from family planning to women's schooling as a more cost-effective way of reducing fertility (Cleland, Bernstein et al. 2006, Joshi 2011, Molyneaux, Gertler 2000). A main reason is the lack of evidence on the effectiveness of family planning programmes in reducing fertility, and this is in turn largely due to the difficulties involved in empirically assessing programme effectiveness (Pörtner, Beegle et al. 2011).

This is also true for Indonesia, which established a national family planning programme in 1969. While there is no questioning that total fertility has declined rapidly in Indonesia since the early 1970s: mean total fertility decreased by 57 percent, from six births in 1971 to 2.5 births in 1999 (Figure 1), there is no consensus over the relative contribution of family planning to the observed decline.

The success of the Indonesian family planning programme as reported in the late 1970s was regarded with great skepticism by many (Cairns Sinquefield, Sungkono 1979). More recent quantitative evidence on the effectiveness of the family planning programme in reducing fertility in Indonesia is not unambiguous either; some studies find no significant effect, and others a negative, significant effect but that varies considerably in size (Angeles, Guilkey et al. 2005, Barnum 1988, Molyneaux, Gertler 2000, Pitt, Rosenzweig et al. 1993). Moreover, the effect of family planning on total fertility in Indonesia has been assessed at the province and district levels rather than at the individual level except for one study (Angeles, Guilkey et al. 2005).

In light of the benefits from lower fertility at the individual level and the large efforts directed towards reducing fertility, it is highly desirable to understand the relative importance of family planning services in lowering fertility.

This chapter empirically examines the effectiveness of family planning in reducing total fertility at the individual level in Indonesia since the early 1970s. Two key proximate determinants of fertility, infant mortality and mother's schooling, are very likely simultaneously determined with fertility. Many studies find an empirical link from infant mortality decline to fertility decline but at the parental level, decisions that raise the risks of infant mortality are also likely to affect fertility and vice-versa. Most existing evidence finds that mother's schooling influences total fertility but it is often treated as exogenous although this may not be the case, as staying on at school might be affected by all alternative activities including starting a family. Given the likely endogeneity of both infant mortality and mother's schooling, this chapter uses instrumental variables to obtain consistent estimates and to capture their causal effects on total fertility.

To examine the effect of family planning on total fertility, a family planning programme exposure variable is constructed that varies across women due to variation in the year of establishment of community-level health posts that provide family planning services (posyandus) and differences in the timing of births within communities. For the estimates to capture the exogenous effect of exposure to family planning services on total fertility the allocation of posyandus should have been random. However, in Indonesia allocation was targeted toward higher fertility areas. To attenuate any bias due to the targeted allocation district specific effects are included in the model. Only one other study has examined the effect of family planning on fertility in Indonesia at the individual level but without incorporating the effect of infant mortality on fertility (Angeles, Guilkey et al. 2005).

Few studies have assessed the causal effect of infant mortality on total fertility due to the difficulty in identifying instruments that are correlated with infant mortality but do not directly affect total fertility (Lee, Schultz 1982). This leaves the option of omitting infant mortality in which case some of the effect picked up by included household and community variables will work through their intermediate effect on infant mortality (Pitt, Rosenzweig et al. 1993, Schultz 1997). However, this chapter uses a public water supply and sanitation programme implemented in a staggered fashion across districts in Indonesia and differences in the timing of births across women to identify exogenous variation in infant mortality that can be used to evaluate the causal effect of infant deaths on total fertility.

Mother's schooling influences fertility decisions but only one study has previously evaluated the causal effect of mother's schooling on fertility in Indonesia using an instrumental variable approach (Breierova, Duflo 2003). This chapter uses a different instrument, the schooling of the fathers of the women who give birth, which it is argued constitutes a plausibly exogenous instrument, to assess the causal effect of mother's schooling on total fertility.

The findings in this chapter strongly suggest that family planning contributed to the fertility decline in Indonesia. Moreover, according to the instrumental variable estimates increases in the number of infant deaths cause higher fertility whereas increases in mother's schooling causes lower fertility.

The rest of the chapter is divided up as follows. Section 1.2 reviews the international and the Indonesian empirical literature most relevant to the analysis in this chapter, and that provides points of comparison for the new results. Next, Section 1.3 provides an introduction to the factors that influence fertility decisions in a simple conceptual framework and discusses the main determinants of fertility. In Section 1.4, the Indonesian national family planning programme and the water supply and sanitation programme used to construct the instrument variable for infant mortality are each described. Section 1.5 discusses the sample used for the analysis and describes the data and variables included in the model. Section 1.6 then considers the issues involved in consistently estimating the effect of family planning, infant mortality, and mother's schooling on fertility, and outlines the empirical models and estimation methods. In

Section 1.7, the results are presented and related to existing evidence. Finally, Section 1.8 concludes.

1.2 Literature Review

In many developing countries total fertility has decreased dramatically since the 1960s (United Nations 2002). The desire to understand what has driven this decline has generated an extensive literature on the determinants of fertility (Schultz 1997, Birdsall 1998). Whereas there is broad consensus that certain variables such as infant mortality and mother's schooling affect fertility, there is disagreement over the importance of others such as family planning programmes (Das Gupta, Boongarts et al. 2011, Joshi 2011, Pritchett 1994, Rosenzweig, Schultz 2008, Schultz 1997, Schultz 1983). This is largely due to the difficulty of measuring family planning programme effectiveness and the subsequent lack of empirical evidence (Pörtner, Beegle et al. 2011).

This section provides an overview of the existing international and Indonesian microeconomic fertility literature with a focus on the role of family planning programmes, infant mortality, and mother's schooling, the variables of main interest in this chapter. The studies reviewed use either cross-sectional or panel data where the woman is the unit of analysis.[1] There are some exceptions for Indonesian studies that use province or district level data given that the number of studies on Indonesia is relatively small. Different approaches are employed to address the issues involved such as endogenous placement of family planning programmes and the likely endogeneity of infant deaths and of schooling, when attempting to consistently estimate the effect of family planning on fertility. Below, studies that cover these issues and that are closest in approach to the analysis in this chapter are discussed in some detail.

1.2.1 International Evidence

An early study combines household survey data with district data for 1968-1971 to examine the effect of family planning and health programmes, and of water and schooling infrastructure on fertility, child mortality, and child schooling in India (Rosenzweig 1982). For a sample of 1,137 rural households simultaneous equations are estimated for fertility, child mortality, and children's schooling to examine the effects of

---

[1] Cross-country studies are not included. For an example see (Angeles, Dietrich et al. 2001).

the public interventions while addressing the endogeneity of fertility, child mortality, and child schooling.

The outcome variable of main interest for the purposes of this chapter is fertility, defined as the number of children born during the 2.5 years prior to the survey. The measures of family planning, health infrastructure, and water supply are constructed as the fraction of villages in each district that has each type of public programme or infrastructure. Interaction terms for parental schooling and family planning exposure are also included to examine whether family planning programmes are more effective when parents have more schooling.

The results suggest that exposure to family planning significantly reduces fertility. However, neither mother's nor father's schooling has any significant effect, and nor are family planning programmes significantly more effective for more schooled parents. Conditional on mother's schooling and age, father's schooling, whether a family lives on a farm or not, and district health, schooling, and water infrastructure, a doubling of villages with a family planning clinic in a district is estimated to lower fertility by 13 percent.

These results rely on the assumption that the family planning programme was randomly allocated with respect to household variables that affect fertility. However, if family planning programmes are not allocated randomly but are targeted to local needs or characteristics ordinary least squares estimates will not only capture the effects of the programmes but also the effect of fertility (and proximate determinants) on programme placement (Molyneaux, Gertler 2000, Pitt, Rosenzweig et al. 1993).

A study of the effect of family planning on fertility in Tanzania directly incorporates the possibility of non-random placement of family planning services (Angeles, Guilkey et al. 1998). The sample consists of 5,077 women in 233 rural villages in Tanzania and fertility is measured by a dummy variable taking the value one if a woman had a birth in each year from age 12, which yields 61,293 woman-year observations.

Contemporaneous access to family planning is captured by nine dummy variables that indicate whether in a given year a woman had access to a health center, hospital or dispensary that provided family planning and the distance to this facility: 5 kilometres, 6-10 kilometres, or 11-30 kilometres. The model also includes age dummies for one of each of five groups (12-14, 15-19, 20-24, 25-29, and 30-34); schooling defined as no schooling, and 1-6 years, 7 years, and more than 8 years of schooling; availability of family planning when a woman was 12 years old; community level child mortality; and time effect.

The family planning facility allocation equations for hospitals, health centres, and dispensaries primarily is intended to allow for consistent estimation of the impact of family planning access on fertility. The covariates included are: community child-mortality rates; the availability of each type of health facility within 30 kilometres; national and regional public spending on health; time effects; and district population as a share of the total population. Few of these variables are individually significant and a joint test of the significance of the identifying variables (national public health spending, regional public health spending, and population shares) has a p-value of 0.063 implying that identification may be weak. Moreover, each of the identifying variables may not only affect programme allocation but also fertility directly.

Three methods are used to estimate the effect of family planning access on fertility: a logit model that does not account for potentially endogenous programme placement and a fixed effects model where community dummies control for unobservables, and a random effects model that is estimated simultaneously with the facility allocation equation.

There are large differences in the estimates for the logit model that does not account for endogenous programme placement and the fixed effects and random effects models that do and only the results for the latter two are reviewed. Women in all age groups have significantly higher probability of a birth than women in the 12-14 years old group as would be expected. All levels of schooling are associated with lower birth probabilities compared to women with no schooling, and the effect is largest for women with eight years or more of schooling. When it comes to access to family planning, hospitals do not significantly impact the probability of a birth, nor does the presence of a hospital, health centre, or dispensary at age 12 suggesting that early exposure to family planning does not affect schooling decisions. By contrast, the contemporaneous measure of having a health center with family planning services within 5 kilometres significantly reduces the probability of a birth. Simulations are also conducted and these indicate that that number of children ever born is 11-13 percent lower for women with contemporaneous access to a health centre with family planning within 5 kilometres relative to women with no such access. The effect of having a facility further away than 5 kilometres are never significant implying that distance matters.

Another recent study of the effect of family planning on fertility in Ethiopia also explicitly models the allocation of family planning services, and examines whether family planning is more effective in reducing fertility for women with certain levels of schooling (Pörtner, Beegle et al. 2011). The main sample consists of 1,326 ever married or partnered women with no formal schooling and the additional sample of 2,051 women with different levels of schooling.

The measure of family planning exposure is based on whether a woman resides in a community that has a health facility that provides family planning services or there is a facility within 40 kilometres, together with the year family planning services were introduced. For this measure of family planning exposure, if some facilities are not used due to remoteness the effect of family planning on fertility will be underestimated. Moreover, some communities may be assigned as only having had family planning services for a relatively short period although a neighboring community has services for a longer time since there is only data on access to the family planning programme that is closest, and this would underestimate the effect of family planning.

A two-stage instrumental variable approach is used where the first-stage equation captures the programme allocation decision as programme placement is assumed to be endogenous, and the second-stage equation the effect of family planning exposure on the outcome of interest: fertility, which is defined as the number of children ever born.

The identification strategy is based on the idea that areas compete for limited resources and that ordinal rankings can be used to discern between their demands. That is, the ordinal rankings of areas are intended to reflect factors taken into account by agencies responsible for allocating resources but that do not directly influence fertility. The first-stage programme allocation equation includes district area; district average annual rainfall and its square; district altitude and its square; rural/urban location; if there is a market; and road accessibility. The excluded instruments assumed to affect fertility only through programme placement are the relative rankings at the district level by population size; share of the population that is urban; and share of adults with different levels of schooling, and at the community level, population size. For instance, districts that are relatively more urbanized are more likely to have a family planning programme.

If the rankings are correlated with the allocation of other programmes, for example, school construction, which also affects fertility, the effect of family planning access will be overestimated as the expectation is for more schooled women to have fewer children.

The F-test for the excluded instruments being jointly equal to zero is 6.14, which is somewhat low with possible implications for consistency and bias due to weak correlation between the excluded instruments and programme placement (Bound, Jaeger et al. 1995, Staiger, Stock 1997).

The model is estimated using ordinary least squares (OLS) and instrumental variables (IV). This may pose a problem as the dependent variable, children ever born, is a count variable. According to the OLS estimates access to family planning significantly lowers the number of children ever born by 0.7 children, and adjusting for endogenous programme placement the IV estimates suggest that the effect is even larger at 0.9 children for women with no schooling. Interacting family planning access with five-year age groups indicates that the effect varies by age. For women less than 20 years old the IV estimates suggest that family planning reduces fertility by 1 child, and for women 40-45 years old by 1.3 children implying a decrease in intervals between births. Including women of different levels of schooling suggests that there is no impact of family planning on fertility for women who have completed first grade or more.

The above approaches are non-experimental. One example of an experimental approach is the evaluation of the Outreach Family Planning and Health Services project (FPMCH) in the rural district Matlab in Bangladesh (Schultz 2008, Schultz, Joshi 2007).[2] The door-to-door FPMCH entailed field workers visiting women of reproductive age roughly every two weeks to provide advice on contraceptive use and contraceptive supplies over the period 1977 to 1996. In 1996, a comprehensive survey in Matlab district interviewed households in 70 villages that had received the FPMCH (treatment villages) and 71 villages that had not (control villages) to assess among other things fertility levels.

Based on a sample of 5,379 married women reduced-form equations are estimated to assess the effects of the FPMCH and other covariates on fertility measured as the number of children ever born. Treated women (those who lived in a treatment village) have significantly lower fertility: the average reduction in fertility for women aged 45-50 years is 1.5 children and for women aged 30-55 years roughly 1.0 child compared to women in the control group. Women with more schooling also have lower fertility, about 0.064 fewer children ever born, although this effect is the same in the treatment

---

[2] "The project seemingly satisfies the definition of a formal experiment, with a well-defined 'treatment' area where services are introduced and a 'comparison' area where such services are absent, but geographical, social, economic, demographic, political and historical conditions are much the same." (Schultz, Joshi 2007)

and control villages suggesting that family planning effectiveness and women's schooling are not substitutes, and that the FPMCH did not affect the fertility differences between relatively more or less schooled women. The results also show that women living in villages in which households are generally wealthier (villages with paved road or a town motor boat) have higher fertility suggesting that wealthier households have more children.

Estimating the reduced-form equation with the fraction of children alive by age 5 as the dependent variable (instead of children ever born), the FPMCH significantly reduced child mortality by 2.5 and 5.6 children per 100 births for women aged 30-35 years and aged 35-40 years respectively. These reductions in child mortality associated with the FPMCH accounts for almost half of the FPMCH related decline in fertility for women aged 45-50 years. Thus, it appears that the maternal and child health component of the FPMCH reduced child mortality, which in turn reduced fertility.

Because family planning programmes are frequently part of maternal and child health programmes or general health programmes in developing countries, the effect of family planning programmes may be overestimated unless infant and child mortality are incorporated into the analysis (Joshi 2011, Schultz 1997). Several studies estimate the effect of infant and child mortality on subsequent fertility without taking family planning into consideration, typically by adding infant or child mortality to reduced form fertility equations (Schultz 1997, Strauss, Thomas 1995, Wolpin 1997).[3] These studies largely fail to address the issue of infant mortality and fertility being simultaneously determined although there are a couple of exceptions (Barnum 1988, Benefo, Schultz 1996).

A study for Colombia using 1973 Census data examines the effect of total child deaths on children ever born through direct estimation rather than indirectly through mortality rates or birth intervals (Olsen 1980).[4] Several sets of regressions are run to assess the bias from using ordinary least squares to enable corrections. The uncorrected estimates from a regression of children ever born on total child deaths imply a replacement rate of more than 1 for both rural and urban women. The estimates

[3] Some studies that examine the effect of infant mortality on fertility use the parity progression ratio (PPR) and the mean closed interval to the next birth (DMCI) (Wolpin 1997). However, these studies will yield inconsistent estimates of the effect of mortality on fertility due to mortality being determined simultaneously with fertility (Lee, Schultz 1982). Given the measure of fertility of interest in this chapter, children ever born, studies that use PPR or DCMI are not reviewed, only those that use the total number of births and deaths in order to facilitate comparisons of results.

[4] Sample size is not reported.

corrected for bias and assuming that child mortality rate varies across women gives a replacement effect of approximately 0.2 to 0.3 births; assuming in addition that the mortality rate is correlated with births, the estimated replacement effects lies between 0.13 and 0.24 births. Moreover, the replacement rates are always higher for urban than for rural women. Another study using the same method for 4,896 women in Korea in 1971 finds that an increase in child mortality is associated with 0.35 to 0.51 more births (Lee, Schultz 1982).

An examination of the effect of total child deaths on total births in Brazil in 1970 uses the correction approach and a minimum distance estimation model for a sample of 30,612 women aged 40-50 years (Mauskopf, Wallace 1984). For the full sample the replacement rate is estimated at 0.6 births but it varies by schooling level and estimated replacement rate is: 0.35 for women with no schooling, 0.6 for those with 1-4 years of schooling, and nearly 1 for women with 5 years or more of schooling.

The impact of child mortality on fertility accounting for the simultaneous determination of these two variables is assessed for women in Côte d'Ivoire and Ghana in the 1980s using an instrumental variable approach (Benefo, Schultz 1996). The sample consists of 1,943 women in Côte d'Ivoire and 2,237 in Ghana who gave birth at least once 5 years or more prior to the survey. Three models are estimated with children ever born used as the measure of fertility. In the first model child mortality is treated as exogenous. The second model treats child mortality as endogenous and instrumental variable estimation is used. The final model adds husband's characteristics and household income. The exogenous covariates included in the models are mother's age, height, and schooling level, household assets, husband's characteristics (assumed to be exogenous), geographic location, ethnicity, religion, distance to market, district mean rainfall per year, and the proportion of households growing tree crops.

The excluded instruments consist of community-level variables: the proportion of households with toilet/latrine and the proportion of households with a protected water source; distance to nearest health clinic and public per capita health spending; prices for staple foods; dummies for the most common community health problem (malaria, diarrhea, or infectious diseases); and a dummy to indicate whether there was a child immunization campaign during the last 5 years. For these instruments to be exogenous they must only affect fertility through child mortality, which is unlikely to be the case for at least the health related variables. Moreover, the instruments are weak: the F-statistic for all excluded instruments being jointly equal to zero is 2.61 for Côte d'Ivoire and 3.06 for Ghana (Bound, Jaeger et al. 1995, Staiger, Stock 1997).

For both countries, child mortality increases fertility significantly when it is assumed to be exogenous (first model); but these effects are small compared to those in other studies. Moreover, there is no significant effect when child mortality is treated as endogenous; arguably this is because of the weak instruments. Household assets are positively and significantly associated with lower fertility in Côte d'Ivoire but not when the husband's characteristics are added. In Ghana, household assets are never significantly related to fertility. Mother's height, intended to capture health status and productivity is positively and significantly related to fertility in Côte d'Ivoire but not in Ghana. Women who have attended middle school and who have attended secondary or higher education each have significantly lower fertility than women with no schooling or some or complete primary schooling, and each additional year of schooling is associated with 0.1 to 0.2 fewer births. The same models are estimated with infant mortality instead of child mortality and the results are similar.

The majority of the fertility literature finds that a woman's schooling (generally 4 years or above) influences fertility (Birdsall 1998). However, although most studies on fertility that include infant or child mortality deal with their endogeneity, it is rare for them to address the potential endogenity of women's schooling although failure to do so may produce misleading results (Angeles, Guilkey et al. 2005, Thomas 1999).

One exception is a study of the impact of parental schooling on fertility in Indonesia based on a sample of 81,549 women (Breierova, Duflo 2003). It uses the large-scale primary school construction programme (SD INPRES) rolled out in Indonesia starting in 1973/74 to construct instrumental variables that capture exogenous variation in schooling. An individual's exposure to the SD INPRES programme is based on her district and year of birth since the programme rollout was staggered over time and across districts.

The instruments for average family schooling are constructed as interactions of year of birth dummies and programme intensity in the district of birth for each wife and husband. The instrument for the average difference in schooling between husbands and wives is captured by combining the husband's year of birth, the intra-couple age difference, and programme intensity in the husband's birth district. Assuming that if there had been no SD INPRES programme fertility patterns would not have been systematically different in districts that received more schools than in those that received fewer, and that the SD INPRES programme only affected fertility through the quantity of schooling, the instrumental variables will capture the impact of schooling on fertility.

A two-stage instrumental variable approach is used to assess the impact of mother's and father's schooling, and the intra-couple difference in schooling on fertility, measured as children ever born. In the first-stage equation of average parental schooling the F-test for the instruments being jointly equal to zero is only 2.1 and in the equation for the intra-couple schooling difference 1.9, which indicates that the instruments are weak with implications for the estimates and may bias them in the same direction as the ordinary least squares estimates (Bound, Jaeger et al. 1995, Staiger, Stock 1997).

Based on the ordinary least squares estimates, average parental schooling is significantly associated with lower fertility (-0.09) and so is the intra-couple difference in schooling (0.028) suggesting that as the husband's schooling rises relative to that of his wife, fertility increases. However, according to the instrumental variable estimates neither parental schooling nor the intra-couple difference in schooling reduces fertility; the former result in particular is in contrast to the existing literature.

1.2.2 Evidence for Indonesia

Numerous studies document the increase in contraceptive use and decline in total fertility coinciding with the rollout of the National Family Planning Programme (NFPP) in Indonesia (Cairns Sinquefield, Sungkono 1979, Chernichovsky, Meesook 1981, Gertler, Molyneaux 1994, Warwick 1986). Below follows a review of the studies on family planning, infant mortality, and mother's schooling and fertility in Indonesia most relevant to the analysis in this chapter.

One study on the effect of mother's schooling and family planning access on fertility in Indonesia finds that failure to address the endogeneity of mother's schooling tends to overestimate the impact of mother's schooling on fertility (Angeles, Guilkey et al. 2005). The data used come from the 1993 Indonesia Family Life Survey (IFLS) and the sample comprises 4,659 women aged 13-51 years with complete birth histories resulting in 113,995 woman-year observations.

The study uses maximum likelihood to estimate four equations simultaneously: children ever born, woman's schooling, husband's schooling, and age at marriage, to control for the endogeneity of preceding decisions on schooling, age at marriage, and number of children. Fertility is modeled for each woman starting at age 10 as an annual logit event for whether a woman had a live birth or not. The birth outcome depends on a woman's age, schooling level, her husband's level of schooling, whether she is married or not, her current number of children, and contemporaneous and past exposure to family planning, and observed community characteristics and allows for time effects; unobserved individual characteristics (e.g., fecundity and family background) that influence birth probabilities over time; and unobserved community factors (e.g., family size norms) that may also affect fertility. The inclusion of the schooling outcome equation in the empirical model allows the examination of the influence of access to family planning services on schooling and, thus, indirect effects of family planning on fertility. Since the data only include information on ever-married women, an equation for the age at first marriage is used to address the possibility of sample selection if women who choose to marry are systematically different from women who do not in terms of fertility decisions.

Prior studies of fertility in Indonesia suggest that placement of family planning services was endogenous (Gertler, Molyneaux 1994, Molyneaux, Gertler 2000) and therefore regional dummy variables are included after which family planning programme placement is treated as exogenous (Angeles, Guilkey et al. 2005).

In the model, nine variables capture past, current, and the degree of access to family planning services for each woman. First, a woman's contemporaneous access to family planning services is measured by a dummy variable for the presence of each of three types of facilities: a puskemas that offer family planning (within 5 kilometres of the community), a posyandu that offer family planning (in the community), and private facilities that offer family planning services (within 5 kilometres of the community). Second, the duration of exposure to the presence of family planning is captured by the years a particular type of facility offering family planning has been available (years available) as a fraction of years available plus 8 years. The final family planning variables is a dummy variable that takes the value one if there was a puskemas in a woman's village when she was age 7. This variable is intended to account for the possibility that knowing about family planning at an early age may influence a woman's schooling decisions.

The endogenous variables children ever born, marital status, and mother's and husband's schooling are measured by dummy variables for each level of schooling: none, primary, junior secondary, senior secondary, and university. In addition to the endogenous variables, rural/urban location, marital status, age, migration status, time effects, and community-level variables: regional gross domestic product and public spending on development, education, family planning, and health are also included in the model.

In the model taking the endogeneity of schooling into account, women with more schooling have higher birth probabilities than those who did not complete primary school whereas in the simple model only women with university education had a higher birth probability than women with less than primary schooling. Increases in a husband's schooling significantly raise the probability of giving birth. When it comes to family planning neither past nor contemporaneous access significantly reduces fertility. However, the three variables measuring the length of time family planning has been available significantly lower fertility, and the long term impact of access to posyandus and private family planning facilities are relatively large at 0.31 and 0.25 fewer births, whereas the impact of access to a puskesmas is very small.

By contrast, other evidence for Indonesia suggests that more female schooling is associated with lower fertility (Pitt, Rosenzweig et al. 1993). The study examines the differences in estimates of the effects of public family planning, health, and schooling programmes when placement of such programmes is not random but is based on unobserved (or unmeasured) characteristics of the target population.

To assess programme effectiveness and differences in estimates when ignoring non-random programme placement and taking it into account household, cross-sectional census, and sub-district (kecamatan) data on the outcomes of interest, regional characteristics that may affect programme allocation, and on public programmes in 1980 and in 1985 are merged and aggregated up to the sub-district level. For the fertility equation the outcome of interest is children ever born for women aged 25-29 years for which 2,904 sub-district observations are available. The sample is also divided into three sub-samples for women with no schooling, with 1-5 years schooling, and with 6 or more years of schooling.

In terms of public programmes, the first model only includes the family planning programme; the second model adds the schooling and health programmes; and the final model also contains all the public programmes but is estimated using a fixed-effects approach that takes advantage of data being available at two points in time. All three models in addition control for mother's schooling, land owned by households, and the proportion of households located in urban areas. For the fixed effects estimates location-specific effects and aggregate time trends are removed but since the estimates are based on two points in time, any location specific time trends cannot be removed.

The results indicate large differences in estimated programme effects depending on whether programme allocation is treated as exogenous or endogenous. For example, under the assumption that family planning programme placement is exogenous, family planning is found to increase fertility (regardless of whether the schooling and health programmes are included in the model), a contradictory result compared to other available evidence. However, the fixed-effects estimates suggest that increased availability of family planning reduces fertility, although the effect is small. Finally, fertility is significantly lower for women with more schooling.

One example of an attempt to control for potential endogeneity of family planning programme allocation in Indonesia similar to the studies discussed above uses information on how resource allocation decisions are made to model placement of family planning services (Molyneaux, Gertler 2000). The identification strategy is based on the observation that province governments allocate resources across all districts subject to a budget constraint and therefore if the resource allocation increases for one district this implies a reduction in resources for another district. Thus, variables that shift contraceptive demand in a district within a province influences the allocation of resources to other districts in the same province but do not directly affect fertility and can therefore be used as instruments.

For the analysis individual data on fertility, proximate determinants, age, and schooling are aggregated to the district level for 286 districts with nine years of quarterly data from the Indonesian Demographic and Health Surveys in 1991 and 1994. This yields a sample of 10,296 observations for the period 1985/86 to 1993/94. Three measures of family planning inputs, also at the district level, are used: public spending on contraceptive supplies per eligible married couple; number of public health clinics that provide family planning services per 1,000 eligible married couples; and number of village contraceptive distribution centres per 1,000 eligible married couples. In addition to the family planning variables the model also includes age, schooling, income and community-level prices and wages.

Three models are estimated where the dependent variable is the quarterly district birth hazard defined as the proportion of women aged 15-49 who gave birth over a three-month period. The first model uses ordinary least squares assuming that all independent variables are exogenous. The second model is estimated by fixed effects allowing for time and district specific effects. One drawback of both models is that if family planning allocation is not random but rather programme inputs are correlated with observed and unobserved determinants of fertility the estimates will be inconsistent. To address the possibility of endogenous programme placement the model is estimated using an instrumental variable-fixed effects approach with the instruments discussed above. In the first-stage equations for contraceptive subsidies, health clinic availability, and village distribution centers the F-statistic for all instruments being jointly equal to zero is 73, 69, and 20 respectively indicating that the instruments are highly relevant.

For all three models, women who have completed secondary school have significantly lower fertility (-0.07 to -1.04), and in the ordinary least squares model women who completed primary school also have significantly lower fertility (-0.06). When it comes to the impact of family planning on fertility the results vary based on whether programme allocation is treated as random or not. Only in the third model that accounts for non-random programme placement are the effects of all three measures of family planning (contraceptive subsidies, health clinics with family planning, and village contraceptive distribution centres) negative although none are significant, which is argued to be the result of high multicollinearity among the three family planning variables. But together, the three coefficients are significant.

However, there may be additional reasons for the results on family planning. For instance, the analysis covers the period 1985/86 to 1993/94 whereas the family planning programme started in 1970 and 1974 in the majority of districts. Also, capturing family access to planning services by an input measure such as contraceptive subsidies may be problematic if not all resources were used for their intended purpose. Moreover, using access of family planning services at the district rather than community level may not capture the connection between availability of family planning services and individual fertility closely enough.

An earlier study of fertility in Indonesia does not account for potential non-random placement of family planning services but does address the endogeneity of infant mortality and the possibility that family planning may influence fertility through reduced infant mortality (Barnum 1988). The sample is a cross-section of Indonesia’s 26 provinces in 1980. Fertility is defined as children ever born to women aged 15-45 and infant mortality as the proportion of infant deaths per total births and both equations use adult illiteracy rates to control for schooling.

To estimate consistently the effect of family planning on fertility, the infant mortality and fertility equations are estimated simultaneously using two-stage least squares. The excluded instrument in the infant mortality equation is the share of households with a toilet or latrine; nutritional status measured by daily calories per capita; public non-hospital recurrent spending per capita; and public hospital recurrent spending per capita. For the fertility equation the excluded instruments are monthly household expenditures per capita; share of population that is urban; and public family planning spending per capita. The first-stage results are not reported so it is not possible to assess whether the instruments are relevant and the exogeneity of some of the excluded instruments is unclear.

The results imply that higher infant mortality is significantly associated with higher fertility; that wealthier households have significantly more children but that households in urban areas have significantly fewer children; and that higher public spending on family planning programmes is significantly related to lower fertility.

However, one potential issue is that the analysis is conducted at the province level, which may be problematic because there is substantial variation in demographic and economic characteristics across districts within provinces, and also within districts in Indonesia.

1.3 Conceptual Framework and Determinants of Fertility

This section begins by briefly outlining the conceptual framework and then discusses how the factors, which are the focus of this chapter: family planning, infant mortality, and mother's schooling influence fertility.

1.3.1 Conceptual Framework

The economic benefits of children include the value of their labour and help they provide at home; support in old-age; and risk insurance. The non-economic benefits include psychic satisfaction and the desire for the family name to live on. Direct costs of children comprise financial costs of child-rearing (for example, cost of food, housing, clothing, and schooling), the opportunity cost of parents' time spent raising their children, and potential social costs for parents that have a different desired family size than prevailing norm (Becker 1960, Chernichovsky, Meesook 1981, Lloyd, Ivanov 1988, Schultz 1997). From an economic perspective fertility is determined by couples' demand for children and so differential fertility rates across couples reflect differences in preferences, resources, and costs and benefits of having children.

The empirical model (Section 1.6) is based on the general fertility model developed by Wolpin (1997) that explicitly incorporates infant/child mortality and which is reproduced below. In the model, a woman can be in one of three states at each point in time, t: $k_t^0 = 1$ if she is not pregnant and not using contraception and zero otherwise; $k_t^1 = 1$ if she is not pregnant but is using contraception and zero otherwise; and $k_t^2 = 1$ if she is pregnant and zero otherwise. The decision each month is whether or not to use contraception. A woman becomes infertile at time $t = \tau + 1$ and therefore $k_t^0 = 1$ for all $t \geq \tau + g$ where g is the gestation period. The monthly contraceptive use and pregnancy history in period t for each woman is:

(1) $K_t = (k_{t-1}^+ \dots k_1^+)$

where $k_t = (k_t^0, k_t^1)$ and $k_t^+ = (k_t, k_t^2)$. It is assumed that a birth occurs, $n_t = 1$, in period t if a woman has been pregnant for the gestation period. A woman's birth history is then $B_t = (n_{t-1} \dots n_1)$ and N_t is the number of children alive in period t. The number of surviving children is given by:

(2) $N_t = \sum_{k=1}^{t} M_t^{t-k}$

where $M_t^{t-k} = 1$ if a child of age $t - k$ is alive at time t and zero otherwise.[5] The probability that a child will die at the start of period t depends on the mother's age at birth and prior birth-spacing, a couple-specific endowment, θ_1, and a child specific endowment, θ_2. A woman's probability of a pregnancy at time t, p_t, depends on her past contraceptive and pregnancy history, which allows for efficiency in contraceptive use. In this framework, a couple maximizes remaining lifetime utility by choosing contraceptive use, k_t, which is costly, and utility each period is state-dependent:

(3) $U_t = k_t^0 U_t^0(N_t, X_t, \omega_t^0) + k_t^1 U_t^1(N_t, X_t, \omega_t^1) + k_t^2 U_t^2(N_t, X_t, \omega_t^2)$

subject to a budget constraint:

(4) $X_t = X_t(Y_t, k_t, M_t, n_t)$

[5] Where $N_0 = 0$, $N_1 = n_1(1 - d_1^0) = M_1^0$, $N_2 = M_1^0(1 - d_2^1) + n_2(1 - d_2^0) = M_2^1 + M_2^0$.

where X_t is a consumption good and ω_t^i is a random preference parameter for each of the three contraceptive use and pregnancy states.

1.3.2 Determinants of Fertility

Family planning programmes aim to reduce fertility with the objective of improving the lives of women and their family members (Schultz 2008). The design of family planning programmes vary and some designs may be more effective in general, and some may work well in some settings but not in others. For the Indonesian National Family Planning Programme (NFPP) the components included: (i) the establishment of an effective system for distributing contraceptives; (ii) training, information dissemination, and outreach activities to promote the use of contraceptives and smaller family size; and (iii) the use of local government and communities to informally convince married couples to use contraceptives (McNicoll, Singarimbun 1983).

The fertility response to exposure to a family planning programme varies with a couple's age, desired family size, current number of children, fecundity, and resources (Schultz 2008). The direct cost of family planning includes the financial cost of contraceptives, which is frequently too high for poor households and travel costs when health centres that provide family planning services are located far away, and these costs may outweigh the cost of having an additional child. Increased availability and subsidization of family planning can by reducing (or eliminating) these costs contribute to lower fertility (Birdsall 1980).

Households, in particular poorer ones, may be uninformed regarding modern birth control methods and of the potential benefits from having fewer children. They may also find it challenging to make optimal decisions when the planning horizon is long (Das Gupta, Boongarts et al. 2011). To reduce the information gap the design of most family planning programmes include methods to disseminate information. For instance, access to radio or television through which family planning information is disseminated can raise contraceptive use (La Ferrara, Chong et al. 2008, Rogers, Vaughan et al. 1999).

Availability of family planning services may also reduce the social cost of having fewer children by contributing to a smaller family size norm among relatives and friends, and within communities (Chernichovsky, Meesook 1981, Das Gupta, Boongarts et al. 2011). The involvement of local communities in the design and operation of family planning programmes may also help change family size norms and speed up the acceptance of modern methods of family planning. This component may have been particularly pertinent in Indonesia where communities were a vital part of the national effort to spread family planning throughout the country (McNicoll, Singarimbun 1983).

Changes in infant and child mortality are associated with changes in fertility as parents alter their behavior in response. There are three main channels through which changes in the probability of infant and child survival influences fertility levels: (i) the physiological effect; (ii) insurance (hoarding) behaviour; and (iii) replacement behavior (Birdsall 1980, Birdsall 1998, Wolpin 1997).

The physiological effect generally works through the inhibiting effect of breastfeeding on fecundity where if an infant that is being breastfed dies this raises the probability of another birth (Lloyd, Ivanov 1988, Wolpin 1997). Insurance behaviour takes place when couples have more children than they otherwise would in anticipation of potential infant and child deaths (Ben-Porath 1980). Finally, replacement behaviour occurs when couples replace a child who has died with an additional birth (Lloyd, Ivanov 1988, Schultz 1978, Wolpin 1997).

The three channels tend to vary in importance depending on the availability of family planning methods. When family planning methods are unavailable the physiological effects tend to dominate; as family planning methods become available but infant and child survival probabilities are still relatively low, insurance behaviour becomes more important; and finally, when family planning methods are available and infant and child survival probabilities are high replacement behaviour becomes increasingly common. The effect of increased infant and child survival probabilities on fertility may be quite different under insurance and replacement behaviour. Under insurance behaviour an increase in survival probabilities can lead to more than a compensatory decline in fertility whereas under replacement behavior it can never result in a fully compensatory reduction in births (Lloyd, Ivanov 1988).[6]

Fertility is generally lower for women with more schooling (Becker 1960, Schultz 1997). There are several reasons why. First, when girls attend school their marriage age and age at first birth tend to rise relative to women who do not attend school (Birdsall 1980, Cleland, Bernstein et al. 2006).[7] Second, women with more schooling are generally better able to obtain and understand information on how to prevent

[6] Since contraceptive methods are never 100 percent safe.
[7] Assuming that women with more schooling do not have shorter inter-birth intervals, which they generally do not (Cleland, Bernstein et al. 2006).

pregnancy; more likely to use modern family planning methods; and more effective users of family planning methods, although a certain minimum level of schooling may be necessary, typically about four years (Birdsall 1980, Cleland 2002, Pörtner, Beegle et al. 2011, Rosenzweig, Schultz 1989).

Third, women are generally the prime care givers and as their schooling levels rise and they face improved employment and earnings opportunities, the opportunity cost of childrearing increases and they may choose to have smaller families (Chernichovsky, Meesook 1981). Fourth, as women's schooling and incomes rise there is less need for children to work either outside or at home, or to have children for old-age security (Lloyd, Ivanov 1988). Fifth, women with more schooling and higher incomes may in addition trade off more investment (e.g., in child health and schooling) in a smaller number of children against a larger number of children in which they invest less (Becker, Lewis 1973).

Any effect of income on fertility will depend on the type of income (e.g., wealth or wages) that changes and the relative income (wealth) level of households. Wealthier households, where wealth takes the form of income from physical assets and land, tend to have more children. This is because increases in the returns to these assets while raising household endowments are unlikely to raise the opportunity cost of raising children (Schultz 1994, Schultz 2005). However, rises in other types of income such as women's wages may lower fertility as the opportunity cost of childrearing increases due to improved labour market and earning opportunities for women (Schultz 2005). However, in cases where women's wages are low, for example, in agriculture, looking after children may not entail a loss of income in which case fertility may not decline even if income rises (Chernichovsky, Meesook 1981).

Moreover, in poorer households children frequently work outside home and if they in addition do not attend school thereby avoiding the cost of schooling, their income may constitute a relatively large share of household income so that parents have more children. Finally, in poorer families children are more likely to be expected to support parents in old age and if the difference between children's and women's wages is relatively small, parents may have more children because the child can make up for the opportunity cost of childrearing later (Chernichovsky, Meesook 1981).

In the next section, the Indonesian National Family Planning Programme (NFPP) used to examine the effect of family planning exposure at the individual level is described. The Water Supply and Sanitation Programme (WSSP) used as an instrument to address the endogeneity of infant mortality is also outlined.

1.4 The Family Planning and Water Supply & Sanitation Programmes

For the analysis of the effect of family planning exposure on total fertility this chapter takes advantage of the Indonesian National Family Planning Programme (NFPP). To examine the impact of infant deaths on total fertility, exposure to the Water Supply and Sanitation Programme (WSSP) rolled out over the period 1973/74-1978/79 is used to identify the impact of infant deaths on fertility. Both theses public programmes are described below.

1.4.1 The National Family Planning Programme

After Independence in 1945 a pronatalist policy was pursued and dissemination of information on family planning measures and contraceptive distribution to the public was prohibited by law (Chernichovsky, Meesook 1981). In 1967, 75 percent of people in Jakarta reported not knowing any method of family planning and modern contraceptive methods were not introduced in Indonesia until the NFPP was established in 1969 (Cairns Sinquefield, Sungkono 1979, McNicoll, Singarimbun 1983).

Under the new government that came into power in 1966 the official population policy was changed and in 1968 the Government of Indonesia (GOI) formally introduced the NFPP with the direct aim of reducing population growth. The NFPP was initially rolled out in Java and Bali provinces in 1970, in 10 additional provinces (Aceh, North Sumatra, West Sumatra, South Sumatra, Lampung, West Nusatenggara, West Kalimantan, South Kalimantan, North Sulawesi, and South Sulawesi) in 1974, and in the remaining 11 provinces (Riau , Jambi, Bengkulu, East Nusatenggara, Central Kalimantan, East Kalimantan, Central Sulawesi, Southeast Sulawesi, Maluku, Irian Jaya, and East Timor) in 1979 (Chernichovsky, Meesook 1981).

The NFPP had three main components: (i) the establishment of a system for distributing contraceptives; (ii) training, information dissemination, and outreach activities to promote the use of contraceptives and promote smaller family size; and (iii) the use of local government and communities to informally convince married couples to use contraceptives (Hayes, Lewis et al. 2003, McNicoll, Singarimbun 1983).

A National Family Planning Coordinating Board (BKKBN) was set up in 1970 to design and coordinate the NFPP (Chernichovsky, Meesook 1981). The BKKBN was responsible for coordinating the activities of the implementing agencies: the Ministry of Health, armed forces, and private organizations (Barnum 1988).

Medical services were health centre based and took the form of health centre staff providing family planning services each week. To reach couples in rural areas, most health centres had field workers for door-to-door canvassing and mobile health teams. Contraceptive supply depots were also established at village-level to enable effective distribution of contraceptives and local community workers were engaged to deliver contraceptive supplies and spread the small-family norm (Cairns Sinquefield, Sungkono 1979, Chernichovsky, Meesook 1981).

By 1976, there were 2,700 health centres providing family planning services, 20,000 village contraceptive supply depots, and 15,000 village-level supply groups (Cairns Sinquefield, Sungkono 1979). By 1979, the number of health centres offering family planning services in Indonesia had reached more than 5,000 (Chernichovsky, Meesook 1981). Contraceptive use increased rapidly alongside the rollout of health centres providing family planning services. In 1971/72, when the NFPP was rolled out in Java and Bali the share of eligible couples using contraceptive methods was only 2.8 percent; in 1974/75 when the NFPP was expanded to Outer Islands I the share rose to 12. 8 percent; in 1979/1980 when the NFPP was extended to Outer Islands II thereby covering all provinces 30.7 percent of eligible couples used contraceptives, and in 1984/85, 15 years after the NFPP was established, 63 percent of eligible couples were using contraceptive methods (Warwick 1986).

The national secular trend in total fertility coinciding with the rollout of the NFPP is shown in Figure 1, and the corresponding declines in provinces where the NFPP started in 1970, 1974, and 1979 respectively are shown in Figure 2, Figure 3, and Figure 4.

The analysis in this chapter assesses the effects of the first (provision of contraceptive) and second (training, information, and outreach) component of the NFPP by constructing a measure of exposure to family planning based on the start year of posyandus: health centres providing family planning services, in each community, combined with the timing of births for women who gave birth over the period 1958 to 2000 (Section 1.5). According to this measure, the unconditional mean total fertility for women unexposed to the NFPP is 4.1 births and for fully exposed women much lower at 2.1 births (Table 1 and Table 4).

1.4.2 The Water Supply and Sanitation Programme

In 1970, only 1 percent of the rural and 10 percent of the urban population in Indonesia had access to improved water supply, and sanitation and sewage systems were inadequate with severe consequences for health outcomes including morbidity and mortality rates for waterborne diseases such as cholera, diarrhea, dysentry, and typhoid (Sanchez 1979, United Nations Environment Program 1989, World Resources Institute 1988).[8] To address the problem the GOI allocated significant resources to improve water supply and sanitation systems under the second-five year development plan 1973/74-1978/79 (Repelita II).

The Presidential Instruction (INPRES) water supply and sanitation programme (WSSP) used as an instrument for the number of infant deaths was funded by INPRES development spending. The grants consisted of transfers from central government to local government budgets to support local-level development. Over the period 1973/74-1978/79, INPRES grants for water supply and sanitation accounted for 2.1 percent of total development spending (World Bank 1984), and the funding allocated to the WSSP increased substantially over time from US$1.5 million in Repelita I to USS$63.6 million in Repelita II (Sanchez 1979).[9] The implementation of the WSSP was the responsibility of the Ministry of Health and the Ministry of the Interior. Province and district authorities were in charge of constructing the water supply and sanitation systems whereas local authorities together with communities were in charge of operation and maintenance (Sanchez 1979).

At the end of the second five-year development plan in 1980, the share of the rural population with access to improved water supply had risen from 1 percent to 19 percent and the share of the urban population from 10 percent to 35 percent (United Nations Environment Program 1989, World Resources Institute 1988).[10] Among the women in the sample not exposed to the WSSP, the mean number of infant deaths is 0.3 compared to only 0.12 for women who were exposed to the WSSP (Table 2).

---

[8] The technologies included in the category improved water supply are: household connection, public standpipe, borehole, protected dug well, protected spring, and rainwater collection.

[9] Repelita I covered the period 1968/69-1972/1973.

[10] In 2003, roughly 51 percent of the rural population had access to drinking water and 40 percent to sanitation .

1.5 Sample Characteristics and Data

This chapter uses data from the Indonesia Family Life Survey (IFLS). Below the characteristics of the sample used for the analysis of the effects of family planning exposure, infant deaths, and schooling on total fertility are described. The data, variable construction, and summary statistics are also discussed.

1.5.1 Sample Characteristics

The data used for the analysis come from the Indonesia Family Life Survey, a repeated socioeconomic and health survey run by the RAND Corporation and the Center for Population and Policy Studies of the University of Gadjah Mada in Indonesia. The first three rounds of the survey: IFLS1, IFLS2, and IFLS3 conducted in 1993, 1997 and 2000 respectively are used. The 10,435 households consisting of 43,649 individuals interviewed in IFLS3 live in 13 of Indonesia's 26 provinces, and represent approximately 83 percent of the Indonesian population in 1993 (Strauss, Beegle et al. 2004). The IFLS surveys 13 provinces: Sumatra Utara, Sumatra Barat, Sumatra Selantan, Lampung, DKI Jakarta, Java Barat, Java Tengah, DI Yogyakarta, Java Timur, Bali, Nusa Tenggara Barat, Kalimantan Selantan, and Sulawesi Selantan, which are the most densely populated areas in Indonesia (Annex figure 1).

The unit of analysis is ever-married women aged 15-49 years at the time of each IFLS round for whom all births occurred during the period 1958 and 2000. This reduces the sample of 8,270 interviewed ever-married women to 7,120 observations, which means that the results may not apply to women who gave birth prior to 1958. After removing observations for which all variables required for the analysis are not available, the sample is reduced to 3,575 observations. Mean total fertility is lower for the women in the analysis sample (2.8) compared to for the 3,545 excluded women (3.2) and the mean number of infant deaths is lower for women in the analysis sample (0.16) than for the excluded women (0.21), and as a result, the findings may not be valid outside the sample. The higher mean total fertility and greater mean number of infant deaths for the excluded women is partly due to them being older as schooling of their fathers is one of the main limiting variables.

An additional issue is that the data only include ever married women. However, since the vast majority of women in Indonesia marry and childbearing takes place almost exclusively within marriage the sample is still closely aligned with the underlying target population (Cairns Sinquefield, Mason 2001, Sungkono 1979, BPS

1992).[11]

Another potential issue is that the sample excludes women who never chose to give birth so that there may be sample selection bias if the relationship between the explanatory variables of interest and fertility decisions for women who never choose to give birth is significantly different from that for women who choose to give birth. However, most women in Indonesia give birth, for instance, only six percent of women aged 45-49 years had never given birth in 1990, implying that any such bias would be relatively small (Mason 2001).

The district level data on the public water supply and sanitation programme for the period 1973-1978 and the share of children aged 7-12 years in each district in 1971 come from administrative records for 1973-1979 (BAPPENAS 1973, 1974, 1975, 1976, 1978). The IFLS data were merged with the administrative data, a time consuming process as several of the 283 districts (kabupaten) changed name, some multiple times, between 1971 and 2000 when the third round of the IFLS was fielded.

1.5.2 Data

The following sections define and describe the data and variables used in the analysis. The main variables of interest are family planning exposure, the number of infant deaths, and mothers' schooling and which are discussed in more detail. Descriptions and summary statistics for all variables are provided in Table 3 and Table 4.

Total fertility

The outcome of interest is total fertility for each ever-married woman. Total fertility is defined as children ever born for women who gave birth over the period 1958 to 2000. This period is chosen to include both women who were and who were not exposed to family planning and to the water supply and sanitation programme. The birth history for all women was constructed by linking the number of reported births in the 1993, 1997 and 2000 rounds of the IFLS.

In the sample, total fertility ranges from 1 to 10 with a mean of 2.8 and a standard deviation of 1.9 (Table 4). Figure 6 shows the distribution of total fertility for the women in the sample: the most frequent numbers of births is one (29 percent) and two

[11] In 1991, 81 percent of women aged 45-49 years were married, 13.5 percent widowed and 3.8 percent divorced (BPS 1992).

(27 percent). For the majority, 83 percent of women, total fertility lies between 1-4 births and a small share, 6 percent, have given birth 7-10 times.

Table 1 and Table 4 compare total fertility for women who were not, partially, and fully exposed to family planning. The raw data show that mean total fertility is higher for unexposed women: 4.4 births and partially exposed women: 2.8 births relative to fully exposed women: 2.2 births and the differences in total fertility between unexposed and partially exposed women, between unexposed and fully exposed, and between partially and fully exposed women are significantly different. However, these unconditional means do not account for any other differences among women with differential exposure to family planning that may also influence total fertility.

Family planning programme allocation and exposure

The family planning exposure variable takes advantage of the variation in the timing of the introduction of posyandus (community sponsored health post that offers family planning services including contraceptives and education, across communities in Indonesia) and the number of births exposed to the presence of such a post for each woman.

The allocation of family planning activities, captured here by the introduction of posyandus, was based on selected indicators: districts with higher morbidity and fertility and fewer health posts were intended to receive posyandus first (Aziz 1990), indicating that programme allocation was not random with implications for the estimates (Section 1.6). District level data on morbidity and fertility rates in the early 1970s are not available.[12] Instead, to control for the possible effects of the non-random allocation of posyandus, the fraction of children aged 7-12 years in the total population in 1971 for each district (ch71_share) is included in the model as a district-level indicator of fertility. The district mean share of children age 7-12 years is 27 percent with a low of 22 percent and a high of 33 percent (Table 4).

To further address the potential endogeneity of posyandu allocation, district dummies are included for each of the 137 districts in the sample in the analysis after which the allocation of posyandus is treated as exogenous (Angeles, Guilkey et al. 2005). The district dummies also help control for district specific effects such as the level of economic development that can also influence individual fertility. The smallest

[12] Such data are only available at the province level.

number of women in living in any given sample district is 6 and the largest is 118 (DI Yogyakarta). The model does not include a variable to account for whether a woman lives in a rural or urban district because by construction, a district is either (primarily) urban or (primarily) rural so district controls will absorb this potential source of fertility variation. However, the difference in total fertility rates across rural and urban areas in the 1960s and 1970s were minor (Cairns Sinquefield, Sungkono 1979).

To generate the measures of family planning exposure the year a posyandu started operating in a woman's community is combined with the year of birth of the first, second and last child for each woman.[13] Data on the community in which a woman's children were born are not available. It is assumed that the community a woman lived in at the time of the IFLS survey is the community in which she gave birth meaning that the family planning exposure dummies will not capture exposure accurately for women who have moved between or after giving birth. To check the potential magnitude of any such a problem, data on a woman's district at the time of her birth and at the time of the survey are examined. Among the women in the sample, 78 percent lived in their district of birth at the time they participated in the IFLS survey. Among the 22 percent who moved, many women are likely to have moved within or to a neighboring district, as inter-province migration was low in Indonesia at the time (Chernichovsky, Meesook 1981).

Based on the timing of the introduction of a posyandu in the community in which a woman lived at the time of the IFLS survey and relative to the date of each of her births, three dummy variables for no, partial, and full family planning exposure are constructed. Women who had no posyandu present in their community before all their births are assigned a value of one for the no exposure family planning dummy; those in a community where a posyandu was present for their second and subsequent births are assigned a value of one for the partial family planning exposure dummy; and those in a community with a posyandu available during all their births are assigned a value of one for the family planning full exposure variable.

Figure **5** shows exposure to family planning and also to the WSSP (see below) for women in the sample.[14] To the right of the vertical line the dotted area indicate women not exposed to family planning nor the WSSP and the dark shaded women exposed to

[13] The year of birth of a woman's first, second, and last child are constructed using data from the three IFLS rounds.
[14] Partially exposed women are not shown in the figure for clarity's sake.

family planning but not to the WSSP. To the left of the vertical line the white area indicates women exposed both to family planning and the WSSP and the lightly shaded area those not exposed to family planning but exposed to the WSSP. For instance, a woman for whom all births occurred between 1988 and 1993 in a district in which a posyandu was constructed in 1983 and the WSSP was rolled out is captured by the white area.

In the sample, 22 percent of women were not exposed, 19 percent were partially exposed, and 59 percent were fully exposed according to the family planning exposure variable (Table 4). By construction, the birth year of the first child among the unexposed women is earlier than that of the partially and fully exposed ones. The mean year of birth for the first and last child among unexposed women is 1975 and 1984; for partially exposed women 1983 and 1990; and for fully exposed women 1992 and 1995 (Table 1). Unexposed women also tend to have less schooling (4.9 years) than partially (5.9 years) and fully (7.6 years) exposed women due to the massive expansion of primary education in Indonesia during the 1970s and 1980s, which raised enrollment substantially for later cohorts.

Data on community level introduction of posyandus is missing for a number of women. As a robustness check, a second set of family planning exposure dummies is constructed. For these dummies, women who are missing data on the year a posyandu started operating in their community are instead assigned the mean posyandu start year for all communities in their district for which the start year is available. This second set of family planning exposure dummies are only used as a robustness check (Table 11-Table 14).

One advantage of this variable compared to other similar ones used in the literature is that for a woman to count as exposed there must be a posyandu in her community rather than within a certain distance, for example, 5 or 40 kilometres (Angeles, Guilkey et al. 2005, Angeles, Guilkey et al. 1998, Pörtner, Beegle et al. 2011). The disadvantage is that a woman who does not have a posyandu in her own community will be assigned as not exposed to family planning even if she is attending a posyandu located in a neighboring community. Similarly, other family planning activities may have taken place at the same time as posyandus were introduced implying that some women treated as unexposed in the analysis were actually exposed to family planning, or that some women assigned partial exposure were fully exposed. This issue cannot be addressed directly and the effect of family planning on total fertility may therefore be underestimated. However, given that the unexposed women consists primarily of women who gave birth when few posyandus were available in the country and exposed women of those who gave birth when posyandus were becoming commonplace, this is likely to be a relatively minor issue.

If a woman does not accurately recall the year of birth for her first, second, and last child, the years used to construct the family planning exposure variable, measurement error may bias the estimates by assigning the wrong type of exposure (e.g., fully exposed instead of partially exposed) to a woman. Yet, given that the maximum age of a woman at the time of interview is 49 years and many women are younger this should not be a major concern.

Programmes may also have cross-effects (Pitt, Rosenzweig et al. 1993). The availability of family planning services may reduce not only fertility but also infant mortality, for example, if women increase the spacing between births or choose to have their first child at a later age than they otherwise would, and also because family planning services are provided at posyandus (community sponsored health posts). In this case, the effect of family planning exposure will be overestimated. But controlling for infant deaths using an instrument will alleviate any such bias.

Infant deaths

To examine the relationship between infant mortality and total fertility the number of infant deaths for each woman is used. An infant death is defined as the death of a child between ages 0-12 months. The majority of women in the sample, 88 percent did not experience any infant death, 9 percent experienced one infant death, 2 percent two infant deaths, and 1 percent three infants deaths (Table 5). The unconditional differences in the mean number of infant deaths across unexposed and partially exposed women and across unexposed and fully exposed women are significantly different, and unexposed women have a higher mean number of infant deaths than either partially or fully exposed women (Table 4).

Simultaneous to the decline in total fertility in Indonesia has been a substantial decline in infant mortality from 145 deaths per 1,000 live births in 1971 to 46 deaths per 1,000 live births in 1999 (Figure 1). If infant deaths and total fertility are determined simultaneously, i.e., more infant deaths lead to higher total fertility, for instance, due to the wish to replace infants who died, and higher total fertility leads to more infant deaths, for example, because of shorter birth spacing intervals or less time to care for each child, then infant deaths will be endogenous in the analysis of total fertility (Barnum 1988, Birdsall 1980, Schultz 1997). As a result, ordinary least squares estimates of the effect of infant deaths on fertility would be biased upwards. This would also be the case because women who have more births experience more deaths since “their ‘sample’ size is larger” (Wolpin 1997: 503).

Infant deaths will also be endogenous if observables that are correlated with infant deaths also influence total fertility (e.g., schooling) are excluded from the model (Schultz 1978). If there are unobservable factors, for instance, fecundity or taste preferences, which affect both infant deaths and total fertility omitted from the model these may induce spurious correlations between and infant deaths will again be endogenous (Ben-Porath 1980, Schultz 1978).

Measurement error with respect to the number of infant deaths is expected to be minor because arguably most women will recall accurately these events given the relatively small number of deaths in each relevant case. To address the potential endogeneity of infant deaths the water supply and sanitation programme described in Section 1.4 is used to construct an instrument for the number of infant deaths.

Water supply and sanitation programme exposure

Existing evidence shows that access to safe water supplies and adequate sanitation can reduce infant mortality (Gunther, Fink 2011). Under the assumption that access to safe water and sanitation only affects fertility through infant deaths, this chapter uses the WSSP discussed in Section. 1.4 as an instrument for the number of infant deaths.

The instrument takes advantage of the variation in the timing and geographical allocation of district spending on the WSSP. The instrument is a dummy variable that takes the value zero for women for whom all births occurred between 1958 and 1978 because the WSSP started between 1973/74 and 1978/79, and the value one for women for whom all births took place between 1979 and 2000 in a district that participated in the WSSP (Figure **5**). In cases where some but not all births of a woman occurred after the start of the WSSP a value of zero is assigned, i.e., the women are assumed to be unexposed.

In the sample, 78 percent of women were exposed to the WSSP as measured by the dummy variable (Table 4). A breakdown of the data into unexposed (and potentially partially exposed) women and exposed women shows that the mean number of infant deaths is more than twice as high, 0.30, for unexposed women relative to those exposed, 0.11 (Table 2). This simple breakdown does not account for any other factors that also influence the number of infant deaths.

The share exposed to the WSSP is greater for women partially exposed to the family planning programme (75 percent) and even greater for women fully exposed (98 percent) than for unexposed women (29 percent) due to the rollout of both programmes starting in the 1970s although across different communities/districts (Table 4). The validity of the water supply and sanitation instrument is discussed in Section 1.6.

Schooling

Schooling is defined as the number of years of formal schooling attained by a woman. Years of schooling is constructed by combining the reported highest grade and level of schooling completed. The Indonesian school system consists of six years of primary school, three years of junior secondary school, three years of senior secondary school, and higher education/university that ranges from two years and up (Angeles, Guilkey et al. 2005).

Mean years of schooling for women in the sample is 6.7 with a standard deviation of 4.1 years. Schooling ranges from 0 (no education) to 18 years (university education) with schooling heaped around no schooling (11 percent), complete primary school (30 percent), complete junior secondary school (13 percent), complete senior secondary school (17 percent), and university (4 percent) (Figure 7).

Mean schooling is 4.9 years for women unexposed to family planning, 5.9 years for partially exposed women, and 7.6 years for fully exposed women (Table 4). The higher schooling on average for partially and fully exposed women is due to the primary school construction programme (SD INPRES) that began in 1973/74 and the ensuing increase in school enrollment and attainment.

Schooling is another potentially endogenous variable in the analysis of total fertility. It will be endogenous if there is unobserved heterogeneity, for instance, innate ability or preferences, which affects both schooling levels and total fertility (Schultz 2008, Sander 1992). Schooling may also be measured with error when recall periods become longer, which would also bias estimates (McCrary, Royer 2011). To deal with the possible endogeneity of schooling, years of schooling of the fathers to the women who give birth is used as the excluded instrument.

Father's schooling

Fathers' schooling is defined the same way as their daughters' schooling: the formal years of schooling attained. To address the potential endogeneity of women's schooling their father's years of schooling is used to capture exogenous variation in schooling. A larger share of fathers, 28 percent, has no schooling compared to their daughters, 11 percent. Moreover, mean years of schooling for fathers is 4.4, also substantially lower than the mean schooling of their daughters due to the expansion of primary schooling since the 1970s.

Another possible instrumental variable for schooling is provided by the massive primary school construction programme 1973/74 to 1978/79 (Breierova, Duflo 2003). But this instrument cannot be used for the analysis in this chapter because of the timing of the family planning and primary school construction programmes being close in time so that many women exposed to the school construction programme would be too young to have given birth or have given birth multiple times at the time of the interview. The validity of the instrument for schooling is considered in Section 1.6.

Household wealth

To examine the effect of household wealth on total fertility a wealth index is included in the model. The index is constructed based on the ownership or availability of six items: electricity; piped water; separate toilet with a septic tank; refrigerator; electric or gas stove; and television set. Each item is assigned a weight equal to the inverse of the fraction of households that own it so that rarer items receive a larger weight.

The measure of wealth is used instead of household income or consumption, first, because the data available for each woman on household assets, income, and consumption refer to the year of the survey, which frequently does not coincide with the timing of a woman's births. Therefore a measure that tends to change less quickly over time such as the wealth index is preferred to household income or consumption. Second, using the wealth index largely avoids the recall bias and measurement error commonly associated with income and consumption (McKenzie 2005).

The mean wealth index is 5.8 (Table 4). At the two extremes, 7.0 percent of the women in the sample do not own or have access to any of the six items whereas 2.8 percent own or have access to all six. Most households, 92 percent, has access to electricity and has a television set, 62 percent (Table 6). About half, 46 percent of the women live in households that have a toilet with septic tank but only 26 percent have piped water, and even smaller shares, 13 percent and 12 percent respectively have a refrigerator, or an electric or gas stove.

Other public programmes and the fertility trend

If there were any other programmes rolled out at the same time as the family planning programme that also affected total fertility and these are not accounted for the estimates of effect of family planning will be biased. The large-scale primary school construction programme that started in 1973/74 partly overlapped with the rollout of family planning services and to control for the possibility that the school construction programme affected total fertility the fraction of children age 7-12 years enrolled in primary school in each district is included in the model. This variable is used because the school programme was targeted at districts with relatively fewer children enrolled and with relatively fewer schools (Aziz 1990).

To capture the observed secular trend in total fertility in Indonesia since the early 1970s dummy variables for each woman's decade of birth are included in the model (Figure 1). The women in the sample are born in the 1940s (7 percent) up until the 1980s (4 percent) (Table 4) and mean total fertility for the sample declines with each decade of birth as expected.

In the next section the empirical approach used to estimate the total fertility equation and estimation issues such as non-random family planning programme placement and endogenous variables are discussed.

1.6 Empirical Methodology

In this section the endogenous allocation of posyandus; the validity of the instruments for infant deaths and mother's schooling; and choice of estimation methods are discussed. The outcome of interest is total fertility and since this is a count variable, maximum likelihood is used to estimate a simple Poisson model followed by a two-stage instrumental variable approach to account for the likely endogeneity of infant deaths and mother's schooling.

1.6.1 Allocation of Family Planning Services

Under the assumption that the allocation of posyandus was random the model would provide estimates of the effect of family planning exposure on total fertility. However, if the allocation of posyandus is targeted the effect of family planning exposure will be biased (Rosenzweig, Wolpin 1986). The direction of any such bias will depend on the criteria used for the targeting: if family planning services are targeted to areas with the greatest need for family planning (districts with high fertility) the effect will be underestimated but if they are targeted toward areas with the greatest demand for family planning the effect will be overestimated (Angeles, Dietrich et al. 2001, Molyneaux, Gertler 2000, Pitt, Rosenzweig et al. 1993).

In Indonesia, family planning services were targeted toward areas with higher fertility and morbidity (Aziz 1990, Das Gupta, Boongarts et al. 2011, Molyneaux, Gertler 2000), and therefore the effect of family planning exposure will likely be underestimated. Evidence on the effect of family planning on fertility in Indonesia find that after controlling for province specific effects, family planning programme allocation can be treated as exogenous (Angeles, Guilkey et al. 2005). Therefore, to control for the non-random allocation of posyandus providing family planning services the model includes district dummies (Section 1.5). To further address issue, the fraction of children age 7-12 years old in each district in 1971 is included as an indirect measure of initial district fertility.

1.6.2 Validity of the Instruments

For the water supply and sanitation programme to be a valid instrument it must be highly correlated with the endogenous variable (relevance), in this case the number of infant deaths, and it must not be correlated with the error term in the total fertility equation (exogeneity).

The simple correlation between the number of infant deaths and water supply and sanitation exposure is -0.154 and statistically significant at one percent. The instrument relevance can be tested and shows that the excluded instrument is relevant (Section 1.7). For the excluded instrument to be exogenous, water supply and sanitation programme exposure must only affect total fertility through the number of infant deaths. The exogeneity assumption cannot be directly tested since the number of excluded instruments equals the number of endogenous variables. Nevertheless, that the water supply and sanitation programme is exogenous is a plausible assumption, particularly, since exposure is measured at the district level. If there were other public programmes correlated with the water supply and sanitation programme that also affect total fertility the estimates would be biased. However, the family planning programme is directly and the primary school construction programme indirectly controlled for as these programmes may also affect total fertility.

Schooling of fathers is used as an instrument and it is valid if it is highly correlated with their daughters' levels of schooling and only influences total fertility through the schooling of their daughters. The unconditional correlation between schooling and father's schooling is 0.574 and significant at one percent suggesting that father's schooling is a relevant instrument; this is confirmed by testing in Section 1.7.

Similar to the water supply and sanitation instrument, the assumption of exogeneity cannot be tested directly. However, in this case it appears reasonable to assume that father's schooling may not directly affect total fertility other than through the schooling of their daugthers. A father's ability is correlated to his schooling and his ability may in turn be correlated with his daughter's ability. If this is the case, father's schooling would be correlated with the error term in the total fertility equation biasing the estimate of the effect of schooling on total fertility. The reason to expect any such bias to be relatively minor is because during the period when the fathers in the sample were of school-age, income and social class would arguably have been more important in determining years of schooling than ability. This is because this was before the start of the massive public primary school construction programme when only a small share of the total population attended school.

1.6.3 Empirical Total Fertility Model: The Benchmark

The first model is a reduced form model, which estimates the total effect of family planning exposure but does not reveal anything about the channels, for instance changes in contraceptive use, through which family planning influences total fertility. The other main variables of interest: the number of infant deaths and mother's schooling are added in subsequent specifications to examine whether they also affect total fertility. Since the dependent variable, the total fertility of each woman is a count variable a Poisson model is used (Figure 6).[15]. The total fertility model is specified as follows:

$$(5) \quad tfr_{ijk} = c_0 + \alpha_j + \beta_k + \gamma_1 fp_{ijk} + \gamma_2 infant_i + \gamma_3 s_i + \gamma_4 w_i + \gamma_5 ch_j + \gamma_6 en_j + u_{ijk}$$

i=individual
j=district of first birth
k=decade of mother's birth

[15] Based on a likelihood ratio test of overdispersion the Poission model is preferred to a negative binomial model (Long, Freese 2006).

tfr_{ijk} = total fertility
c_0 = constant
α_j = district of first birth
β_k = decade of mother's birth
fp_{ijk} = family planning exposure
$infant_i$=number of infant deaths
s_i=mother's schooling
w_i= household wealth index
ch_j = fraction of district population aged 7-12 years in 1971
en_j=fraction of children enrolled in primary school in 1971
u_{ijk}= unobservables (e.g., fecundity of family size preferences)

1.6.4 Empirical Total Fertility Model: A Two-Stage Approach

To address the potential endogeneity of the number of infant deaths and schooling a two-stage approach with the water supply and sanitation programme and father's schooling as the excluded instruments is used. In the first-stage, the number of infant deaths is regressed on all the exogenous variables including the water supply and sanitation programme, and mother's schooling is regressed on all the exogenous variables including father's schooling to generate the fitted values for the second-stage total fertility regression. In equation 7, γ_1 measures the effect of family planning exposure; γ_2 the effect of infant deaths; and γ_3 the effect of mother's schooling on total fertility. The first-stage equation for infant deaths takes the form:

$$(6) \quad infant_{ijk} = c_0 + \alpha_j + \beta_k + \gamma_1 fp_{ijk} + \gamma_2 wss_{ijk} + \gamma_3 sf_i + \gamma_4 w_i + \gamma_5 ch_j + \gamma_6 en_j + e_{ijk}$$

$infant_{ijk}$= number of infant deaths
wss_{ijk}= water supply and sanitation programme exposure
sf_i=schooling of father of woman who gave birth
e_{ijk}=unobservables (e.g., mother's ability to care for her child and mother's health during pregnancy)
All other variables defined as above

The first-stage schooling equation is given by:

$$(7) \quad s_{ijk} = c_0 + \alpha_j + \beta_k + \gamma_1 fp_{ijk} + \gamma_2 wss_i + \gamma_3 sf_i + \gamma_4 w_i + \gamma_5 ch_j + \gamma_6 en_j + v_{ijk}$$

v_{ijk}=unobservables (e.g., innate ability)
All other variables defined as above

The outcome equation for total fertility is then:

$$(8) \quad tfr_{ijk} = c_0 + \alpha_j + \beta_k + \gamma_1 fp_{ijk} + \gamma_2 \widehat{infant}_i + \gamma_3 \hat{s}_i + \gamma_4 w_i + \gamma_5 ch_j + \gamma_6 en_j + u_{ijk}$$

$\widehat{infant}_i$=predicted number of infant deaths
$\hat{s}_i$=predicted schooling of mother
All other variables defined as above

The identification comes from wss_{ijk} and sf_i in the first-stage. For the estimates to measure the effect of family planning exposure on total fertility it is necessary to assume that family planning programme allocation is not correlated with any unobserved area characteristics, which is plausible given the included district and decade controls.

The schooling equation is estimated by ordinary least squares. However, given the heaping of schooling at the completion of each level of schooling (Section 1.5 and Figure 7), the schooling equation is also estimated using a negative binomial model. The first-stage coefficients have the same signs and significance as the ordinary least squares estimates but differ somewhat in size. However, the second stage coefficients are highly similar in all respects. Given the similarity of the estimates, the fitted values from the first-stage ordinary least squares schooling regression are used in the second-stage total fertility regression.

Family planning exposure and water supply and sanitation exposure both vary at the district level, therefore to allow for correlation in the error structure for women within district all standard errors are clustered at the district level.

The next section presents the findings on the effect of family planning exposure, the number of infant deaths, and mother's schooling on total fertility.

1.7 Results and Discussion

In this section the maximum likelihood estimates of the benchmark Poisson model are presented first. Then the two-stage estimates using the water supply and sanitation programme and father's schooling as instruments for the number of infant deaths and schooling respectively are discussed and compared to the benchmark results.

1.7.1 Results from the Benchmark Model

Table 7 shows the results for the reduced form estimations; the specifications that include the number of infant deaths and schooling; and those that add wealth, for the sample of 3,575 women.

Columns 1-3 contain the reduced form estimates of the overall effect of family planning exposure on total fertility. The first column includes the fraction of children age 7-12 years in each district in 1971 to account for the endogenous allocation of family planning services (posyandus), and the fraction of children enrolled in primary school in each district in 1971 to control for the large-scale school construction programme. The second column adds the district dummy variables to further control for the endogenous family planning allocation and other district specific effects. The third column adds the mother's decade of birth dummies to account for the secular trend in total fertility. In columns 4-6 the endogenous variables: the number of infant deaths and schooling are entered, and finally, in columns 7-9 the wealth index is included.

Given the likely endogeneity of family planning allocation and the observed downward trend in total fertility in Indonesia over time, the specifications that include district and decade dummies are the preferred ones and are discussed in detail (columns 3, 6, and 9).

Starting with family planning exposure it is clear that women who were exposed whether partially or fully, have significantly lower total fertility, and that the effect of family planning exposure on total fertility is relatively large. According to the reduced form estimates, expected total fertility is lower by a factor of 0.80 (21 percent) for partially exposed women and by a factor of 0.76 (24 percent) for fully exposed women holding all other variables constant (column 3). The negative relationship between family planning and fertility is in line with existing individual level studies (Angeles, Dietrich et al. 2001, Angeles, Guilkey et al. 2005, Joshi 2011, Pörtner, Beegle et al. 2011, Rosenzweig, Wolpin 1982).

After including infant deaths and mother's schooling, the effect of family planning is somewhat reduced.[16] Now being partially exposed to family planning reduces mean total fertility by a factor of 0.82 (18 percent) and being fully exposed by a factor of 0.80 (20 percent) with all other variables kept constant (column 6). The proportional effects are smaller than those from the reduced form model as would be expected given the role of infant deaths and mother's schooling in total fertility decisions and this underlines the importance of accounting for these variables to obtain consistent estimates of the effect of family planning. That total fertility of women who were fully exposed to family planning is reduced by more than for women who were partially exposed provides reassurance that the constructed measure of family planning exposure is working as intended.

[16] As a check, the specification in column 9 in Table 7 is estimated but replacing infant deaths by water supply and sanitation programme exposure and the coefficient on wss programme exposure is not statistically significant.

More schooling is also associated with lower total fertility. A one unit increase in schooling reduces expected total fertility by a factor of 0.98 (2.2 percent) holding all other factors constant. The significant negative effect of mother's schooling is found in most of the empirical literature (Angeles, Guilkey et al. 1998, Joshi 2011, Molyneaux, Gertler 2000, Pitt, Rosenzweig et al. 1993,), although there are exceptions (Pörtner, Beegle et al. 2011).

By contrast, women who have experienced more infant deaths have higher total fertility: a one unit increase in the number of infant deaths raises mean total fertility by a factor of 1.24 (24.4 percent) holding the other variables constant, supporting the idea that women have more children to replace children who die. This is similar to other studies both international and for Indonesia (Barnum 1988, Lee, Schultz 1982, Mauskopf, Wallace 1984, Olsen 1980).

The final specification includes wealth, which has a significant but small positive effect on total fertility. A one standard deviation (5.6) increase in wealth (as measured by the wealth index) raises total fertility by a factor of 1.0 (0.4 percent) holding all other variables constant. Other studies also find that wealthier households (those with more assets, not higher income) tend to have higher fertility (Barnum 1988, Schultz, Joshi 2007). The inclusion of wealth does not change the results on family planning exposure, infant deaths, and mother's schooling (column 9).

Including the dummy variables for the decade of birth for each woman reduces the estimated effect of family planning on total fertility. This strongly suggests that failure to control for the downward trend in total fertility over time would overestimate the effect of family planning exposure on total fertility. The inclusion of the district and decade controls also improves the overall fit of the model.

Overall, the evidence presented so far strongly suggests that total fertility is lower for women exposed to family planning and who have more schooling, and that total fertility is higher for women who have experienced more infant deaths and who are wealthier although the latter effect is small.

It remains to address the endogeneity of infant deaths and mother's schooling in order to assess the causal effect of these on total fertility, and to examine the effect of family planning on total fertility after these adjustments.

1.7.2 Instrumental Variable Results

The number of infant deaths and mother's schooling are assumed to be endogenous as discussed above and testing corroborates this.[17] To address the endogeneity of infant deaths and schooling and examine their causal effect on total fertility a two-stage instrumental variable approach is used. In the first-stage, exposure to the water supply and sanitation programme is used as the excluded instrument for the number of infant deaths and father's schooling is used as the excluded instrument for schooling of the women who give birth.

The distribution of the number of infant deaths is a count variable (Table 5) and testing shows that a negative binomial model is preferred to a Poisson model due to overdispersion (Long, Freese 2006).[18] The first-stage schooling equation is estimated by ordinary least squares and the fitted values used in the second-stage total fertility regressions schooling (Section 1.6).

Table 8 shows the first-stage regressions with the number of infant deaths as the dependent variable. Water supply and sanitation programme exposure is always negatively and significantly related to the number of infant deaths and being exposed to the water supply and sanitation programme reduces infant deaths by a factor of 0.65 (35 percent) holding all other variables constant in the preferred specification that includes district and decade dummies (column 3).

The other explanatory variables also have the expected signs: exposure to family planning (partial or full), being a woman with a more schooled father, and being wealthier are all significantly associated with fewer infant deaths.

With respect to effect size, for the family planning programme being partially exposed implies a 35 percent decrease in the number of infant deaths and being fully exposed a 14 percent reduction although the latter is not significant in the preferred specification (column 3). For a one standard deviation (3.8 years) increase in the years

---

[17] In a regression of the number of infant deaths on total fertility, schooling, wealth, exposure to the water supply and sanitation programme, exposure to the family planning programme, and controlling for district and decade of birth for the 3,575 women, higher total fertility is highly significantly associated with more infant deaths. This indicates that total fertility and the number of infant deaths are determined simultaneously. In a similar regression but with schooling as the dependent variable total fertility is highly significantly associated with schooling implying that there is some unobservable factor that determines both schooling and total fertility that is omitted from the model. To assess whether the number of infant deaths is endogenous a generalized Hausman test is used and for schooling a standard Hausman test (Davidson, MacKinnon 1993, Grogger 1990).

[18] There is significant evidence of overdispersion according to the likelihood ratio test G^2=($\ln L_{NBM}$-$\ln L_{Poisson}$) with a Chi^2 distribution adjusted for truncation (Long, Freese 2006).

of schooling of a woman's father, the number of infant deaths is 4 percent lower holding all other variables constant.

Finally, a one standard deviation (5.6) rise in the wealth index reduces the number of infant deaths by nearly 4 percent. In terms of effect size, the most important determinants of the number of infant deaths are therefore exposure to the water supply and sanitation programme and being partially exposed to the family planning programme.

As these are negative binomial regressions, to assess the strength of the excluded instrument a likelihood ratio test is used instead of the standard F-statistic. According to the test the instrument is highly relevant in all three specifications indicating there is no reason to expect bias due to weak correlation between the water supply and sanitation programme and the number of infant deaths (Bound, Jaeger et al. 1995, Staiger, Stock 1997).[19]

The estimates from the first-stage schooling equation are shown in Table 9. Schooling of fathers to the women who give birth is always positively and significantly associated with schooling; one standard deviation increase (3.8 years) in father's schooling implies a 1.4 year increase in daughter's schooling controlling for other factors. The F-statistic is 28.0 in the preferred specification indicating that the instruments are highly relevant (column 3) (Bound, Jaeger et al. 1995, Staiger, Stock 1997).

There is no significant relationship between being partially exposed to the family planning programme and schooling. But full family planning exposure and schooling are positively and significantly related although this effect is relatively small at an additional 0.62 years of schooling. Being exposed to the water supply and sanitation programme is also associated with significantly more schooling and women from wealthier households have significantly more schooling as expected. A one standard deviation increase in the wealth index (5.6) implies an additional 1.3 years of schooling holding other factors constant. Such a wealth increase would be equivalent, for example, to a woman who currently owns none of the six assets included in the wealth index to acquire piped water and a toilet, or to acquire a refrigerator, a large change in wealth.

[19] The model is run with and without the water supply and sanitation programme variable to generate a likelihood ratio test of the significance of the instrument.

Table 10 presents the second-stage estimates of the effect of family planning exposure, infant deaths, and mother's schooling on total fertility obtained by estimating a Poisson model. The standard errors are bootstrapped to account for the two-step estimation process. The results are similar in terms of direction and significance to those from the benchmark Poisson model discussed above (Table 7). Total fertility is significantly lower for women who: were exposed to the family planning programme; have experienced fewer infant deaths; have more schooling; and are less wealthy.[20]

Focusing on the preferred specification (column 3) that includes district and decade controls, expected total fertility of women, both those partially and those fully exposed to family planning is lower by a factor of 0.84 (16 percent) holding all other factors constant. For partially exposed women the effect size is similar to that in the benchmark model, however, for fully exposed women the effect is smaller in proportional terms. Together the benchmark and instrumental variable results strongly suggest that the introduction and expansion of family planning services in Indonesia contributed to lower total fertility.

More infant deaths lead to higher total fertility, arguably as women replace children they have lost. The effect is relatively large: a one unit increase in the number of infant deaths increases total fertility by a factor of 1.15 (15 percent) this is compared to the benchmark model (Table 7) where a one unit increase in the number of infant deaths is associated with an increase in total fertility by a factor of 1.25 (25 percent) holding all other factors constant. This indicates that failure to address the endogeneity of infant deaths would bias the estimates upward.

More schooling causes lower total fertility. A one unit increase in mother's schooling reduces total fertility by a factor of 0.97 (3.0 percent), which is larger than the effect in the benchmark model where one additional year of schooling is associated with a 2.4 percent reduction in expected total fertility. The effect of mother's schooling is much smaller than that of both family planning exposure and the number of infant deaths.

Similar to the results from the benchmark model, being wealthier is not significantly associated with total fertility in the preferred specification (column 3) although it is in

[20] The regressions are also run restricting the sample to women who are 35+ and therefore should have complete or nearly complete fertility histories. The results are similar to those for the main sample indicating that censored fertility histories do not bias the findings from the main sample.

the model that does not control for decade of birth (column 2). This is arguably due to the positive trend in income and wealth over the time period under examination.

The share of children aged 7-12 years in each district is only significant in the model that does not account for district and decade of birth (column 1) suggesting that the district dummies provide better controls for non-random placement of the family planning programme. Likewise, district level primary school enrollment in 1971 is only significant in the specification that excludes district and decade dummies.

To examine the robustness of the results to a change in the sample the same regressions are run but using an adjusted family planning exposure variable that allows the sample to be increased to 4,809 women. For this measure of family planning exposure, women who were previously omitted from the analysis because data on which year a posyandu started operating in their community is missing are assigned as unexposed or exposed based on the mean posyandu start year in their district. This is clearly a less accurate measure of family planning exposure but serves to provide a robustness check.

Comparing the estimates for the benchmark model (Table 7) and the corresponding model for the larger sample (Table 11) the findings are comparable. Being partially exposed to family planning reduces total fertility by a factor of 0.77 (23 percent) and fully exposed by a factor of 0.81 (19 percent) using the alternative measure of family planning compared to a factor of 0.83 (18 percent) and 0.80 (20 percent) respectively for the benchmark model. As expected the difference in results is larger for partially exposed women who are those affected by the adjustment of the variable definition. Having experienced more infant deaths and having less schooling are each significantly associated with lower total fertility and the effect sizes are nearly the same.

In the first-stage regressions of infant deaths (Table 12), the results are similar to those for the main sample (Table 8). However, in the preferred specification that includes decade and district of birth controls (Table 12, column 3), the water supply and sanitation variable is only significantly related to the number of infant deaths at the 11.5 percent level but the likelihood ratio test strongly suggests that the water supply and sanitation programme variable is a relevant instrument.[19] In the first-stage regressions of schooling all results remain practically the same and the instruments are highly relevant (F-statistic of 35) (Table 13).

Table 14 shows the second-stage results for total fertility. The results of family planning exposure (partial and full) are nearly identical to those for the main sample

(Table 10) with total fertility of exposed women being significantly lower. As before, more schooling causes lower total fertility, and having experienced more infant deaths leads to higher total fertility.

Overall, the results provide strong support that the family planning programme in Indonesia contributed to lower total fertility and evidence of a causal link from higher infant mortality to higher total fertility and from more schooling to lower total fertility.

1.8 Conclusion

Indonesia established a national family planning programme in 1969 and total fertility has declined rapidly since but there is little consensus over the relative contribution of family planning to the observed decline. This chapter empirically examines the effectiveness of family planning in reducing total fertility at the individual level in Indonesia and the findings strongly suggest that family planning contributed to the fertility decline. This main result compares to that of a recent study of the effect of family planning of fertility in Ethiopia that uses a similar approach to capture exposure to family planning and finds that it reduces fertility (Pörtner, Beegle et al. 2011).

This chapter also assesses the causal effect of infant mortality on total fertility, which few studies have done because of the difficulty in identifying suitable instruments. However, this chapter takes advantage of a public water supply and sanitation programme implemented in staggered fashion across districts in Indonesia to capture exogenous variation in infant mortality. This instrument is chosen based on the idea that increased access to clear water and sanitation increases the probability of infant survival. The results strongly indicate that increases in the number of infant deaths causes higher total fertility.

The causal effect of mother’s schooling on total fertility has previously been evaluated by one other study for Indonesia (Breierova, Duflo 2003). This chapter uses an alternative instrument, the schooling of fathers of the women who give birth, to capture exogenous variation in schooling and finds that more schooling causes lower total fertility.

Although this chapter shows that family planning contributes to lower fertility together with reductions in infant deaths and improvements in women’s schooling, it does not reveal the mechanisms through which this occurs. Therefore, additional research is required to understand what the underlying mechanisms are and in what settings and forms family planning programmes are most effective in reducing fertility.

Table 1. Mean total fertility by family planning exposure

	Family planning		
	not exposed	partially exposed	fully exposed
mean total fertility[1]	4.4	2.8	2.2
mean schooling	4.9	5.9	7.6
mean birth year first child	1975	1983	1992
mean birth year last child	1984	1990	1995
observations	814	669	2092

Note: 1. Total fertility defined as children ever born for each woman.

Table 2. Mean infant deaths by WSSP exposure

	Water supply and sanitation programme	
	not exposed[1]	fully exposed
mean number of infant deaths	0.30	0.12
observations	779	2796

Note: 1. This group also includes women who were potentially partially exposed to the water supply and sanitation programme.

Table 3. Variable description

Variable	Description
decade	dummy variable for decade of birth for each woman
district	dummy variable for first child's district of birth
FP exposure 2, full	dummy variable that takes value 1 if woman fully exposed to family planning, 0 otherwise, (district mean posyandu start year for women missing community level posyandu start year)
FP exposure 2, partial	dummy variable that takes value 1 if woman partially exposed to family planning, 0 otherwise (district mean posyandu start year used for women missing community level posyandu start year)
FP exposure, full	dummy variable that takes value 1 if woman fully exposed to family planning, 0 otherwise (community level posyandu start year)
FP exposure, partial	dummy variable that takes value 1 if woman partially exposed to family planning, 0 otherwise (community level posyandu start year)
infant deaths	number of children who died between ages 0-12 months for each mother
proportion enrolled in 1971	fraction of children enrolled in primary school in 1971 in each district
proportion of children in 1971	fraction of population aged 7-12 years in 1971 in each district
schooling	years of schooling for each mother
father's schooling	father's years of schooling for each woman who gave birth
total fertility	children ever born for each woman
wealth	based on household access to or ownership of each of six components: electricity; piped water; separate toilet with a septic tank; refrigerator; electric or gas stove; and television set and each component assigned a weight equal to the inverse of the fraction of household that have that particular component
WSS pgm exposure	water supply and sanitation programme exposure for each woman; 0 if not exposed or potentially partially exposed, 1 if fully exposed

Table 4. Sample summary statistics

	Full sample		Family planning					
			Unexposed		Partially exposed		Fully exposed	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
father's schooling	4.4	3.8	3.2	3.6	4.0	3.7	5.1	3.7
FP exposure, full	0.59	0.49	0	0	0	0	1	0
FP exposure, partial	0.19	0.39	0	0	1	0	0	0
infant deaths	0.16	0.48	0.30	0.65	0.14	0.44	0.11	0.39
proportion enrolled in 1971	0.16	0.06	0.17	0.08	0.16	0.06	0.16	0.06
proportion of children in 1971	0.27	0.02	0.27	0.02	0.27	0.02	0.27	0.02
schooling	6.7	4.1	4.9	3.9	5.9	4.1	7.6	3.9
total fertility	2.8	1.9	4.4	2.0	2.8	1.7	2.2	1.5
wealth	5.8	5.6	6.3	6.1	5.6	5.4	5.8	5.5
WSS pgm exposure	0.78	0.41	0.29	0.46	0.75	0.43	0.98	0.14
decade 1940	0.07	0.26	0.23	0.42	0.06	0.24	0.01	0.11
decade 1950	0.21	0.41	0.49	0.50	0.27	0.44	0.09	0.28
decade 1960	0.35	0.48	0.24	0.43	0.46	0.50	0.35	0.48
decade 1970	0.33	0.47	0.04	0.19	0.19	0.39	0.48	0.50
decade 1980	0.04	0.20	0.00	0.00	0.01	0.12	0.07	0.25
Observations	3575		814		669		2092	

Table 5. Frequency of infant deaths

Number of infant deaths	Observations	Percent
0	3138	88
1	331	9
2	77	2
3	29	1
	3,575	100

Table 6. Share of households with each wealth component

Wealth component	Percent
electricity	92
television set	62
toilet with septic tank	46
piped water	26
refrigerator	13
electric or gas stove	12

Table 7. Benchmark total fertility regressions (main sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Poisson	tfr	tfr	tfr	tfr	tfr	tfr	tfr	tfr	tfr
FP exposure, partial	-0.4368***	-0.4414***	-0.2307***	-0.3526***	-0.3528***	-0.1947***	-0.3379***	-0.3398***	-0.1940***
	(0.0330)	(0.0329)	(0.0288)	(0.0317)	(0.0321)	(0.0278)	(0.0314)	(0.0321)	(0.0277)
FP exposure, full	-0.7113***	-0.7767***	-0.2791***	-0.5655***	-0.6122***	-0.2286***	-0.5385***	-0.5844***	-0.2270***
	(0.0239)	(0.0270)	(0.0304)	(0.0266)	(0.0295)	(0.0297)	(0.0265)	(0.0301)	(0.0297)
infant deaths				0.2686***	0.2589***	0.2184***	0.2713***	0.2615***	0.2196***
				(0.0147)	(0.0174)	(0.0141)	(0.0148)	(0.0170)	(0.0141)
schooling				-0.0293***	-0.0355***	-0.0226***	-0.0376***	-0.0421***	-0.0246***
				(0.0033)	(0.0037)	(0.0031)	(0.0034)	(0.0039)	(0.0031)
wealth							0.0128***	0.0138***	0.0038
							(0.0021)	(0.0025)	(0.0025)
proportion of children in 1971	Y	Y	Y	Y	Y	Y	Y	Y	Y
proportion enrolled in 1971	Y	Y	Y	Y	Y	Y	Y	Y	Y
district		Y	Y		Y	Y		Y	Y
decade			Y			Y			Y
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pseudo R-squared	0.070	0.092	0.140	0.102	0.122	0.157	0.105	0.124	0.157
Observations	3575	3575	3575	3575	3575	3575	3575	3575	3575

Robust standard errors in parentheses. Standard errors clustered at district level. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant.

Table 8. First-stage infant deaths regressions (main sample)

Negative binomial	(1) infant deaths	(2) infant deaths	(3) infant deaths
FP exposure, partial	-0.5796***	-0.5406***	-0.4231***
	(0.1721)	(0.1624)	(0.1605)
FP exposure, full	-0.6035***	-0.5455***	-0.1389
	(0.1637)	(0.1640)	(0.1799)
WSS pgm exposure	-0.4615***	-0.6168***	-0.4317**
	(0.1508)	(0.1530)	(0.2117)
father's schooling	-0.0618***	-0.0590***	-0.0425***
	(0.0162)	(0.0151)	(0.0152)
wealth	-0.0481***	-0.0368**	-0.0439***
	(0.0132)	(0.0144)	(0.0141)
proportion of children in 1971	Y	Y	Y
proportion enrolled in 1971	Y	Y	Y
district		Y	Y
decade			Y
Prob > chi2	0.000	0.000	0.000
Pseudo R-squared	0.046	0.120	0.138
Observations	3575	3575	3575

Robust standard errors in parentheses. Standard errors clustered at district level.
***p<0.01, **p<0.05, *p<0.1. All regressions include a constant.

Table 9. First-stage schooling regressions (main sample)

	(1)	(2)	(3)
Ordinary least squares	schooling	schooling	schooling
FP exposure, partial	0.2110	0.2656	0.0827
	(0.1790)	(0.1762)	(0.1689)
FP exposure, full	1.1525***	1.2655***	0.6219***
	(0.1926)	(0.1972)	(0.2038)
WSS pgm exposure	1.3175***	1.4347***	0.8574***
	(0.1808)	(0.1899)	(0.2314)
father's schooling	0.4427***	0.3809***	0.3578***
	(0.0208)	(0.0195)	(0.0190)
wealth	0.2251***	0.2125***	0.2294***
	(0.0114)	(0.0116)	(0.0115)
proportion of children in 1971	Y	Y	Y
proportion enrolled in 1971	Y	Y	Y
district		Y	Y
decade			Y
F test of excluded instruments (F statistic/P-value)	448.58/0.000	26.97/0.000	28.02/0.000
R-squared	0.468	0.529	0.546
Observations	3575	3575	3575

Robust standard errors in parentheses. Standard errors clustered at district level.
***p<0.01, **p<0.05, *p<0.1. All regressions include a constant.

Table 10. Second-stage total fertility regressions (main sample)

Poisson	(1) tfr	(2) tfr	(3) tfr
FP exposure, partial	-0.1879***	-0.2854***	-0.1871***
	(0.0671)	(0.0331)	(0.0272)
FP exposure, full	-0.3774***	-0.4771***	-0.2151***
	(0.0621)	(0.0358)	(0.0303)
infant deaths	1.1610***	0.3561***	0.2431***
	(0.3631)	(0.1197)	(0.0965)
schooling	-0.0317***	-0.0702***	-0.0318***
	(0.0115)	(0.0080)	(0.0075)
wealth	0.0180***	0.0211***	0.0051
	(0.0034)	(0.0027)	(0.0029)
proportion of children in 1971	Y	Y	Y
proportion enrolled in 1971	Y	Y	Y
district		Y	Y
decade			Y
Prob > chi2	0.000	0.000	0.000
Pseudo R-squared	0.079	0.101	0.142
Observations	3575	3575	3575

Bootstrapped standard errors in parentheses. Standard errors clustered at district level. ***p<0.01, **p<0.05, *p<0.1. All regressions include a constant. Infant deaths and schooling are instrumented.

Table 11. Benchmark total fertility regressions (large sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Poisson	tfr	tfr	tfr	tfr	tfr	tfr	tfr	tfr	tfr
FP exposure 2, partial	-0.4368***	-0.4407***	-0.2349***	-0.3617***	-0.3624***	-0.2041***	-0.3427***	-0.3452***	-0.2026***
	(0.0286)	(0.0277)	(0.0249)	(0.0277)	(0.0271)	(0.0242)	(0.0279)	(0.0273)	(0.0242)
FP exposure 2, full	-0.7557***	-0.8060***	-0.3113***	-0.6117***	-0.6551***	-0.2587***	-0.5817***	-0.6231***	-0.2568***
	(0.0220)	(0.0226)	(0.0295)	(0.0241)	(0.0247)	(0.0274)	(0.0246)	(0.0257)	(0.0276)
infant deaths				0.2730***	0.2615***	0.2188***	0.2770***	0.2650***	0.2205***
				(0.0142)	(0.0160)	(0.0132)	(0.0141)	(0.0155)	(0.0131)
schooling				-0.0263***	-0.0297***	-0.0206***	-0.0354***	-0.0378***	-0.0232***
				(0.0028)	(0.0032)	(0.0027)	(0.0030)	(0.0034)	(0.0027)
wealth							0.0126***	0.0139***	0.0041**
							(0.0017)	(0.0020)	(0.0020)
proportion of children in 1971	Y	Y	Y	Y	Y	Y	Y	Y	Y
proportion enrolled in 1971	Y	Y	Y	Y	Y	Y	Y	Y	Y
district		Y	Y		Y	Y		Y	Y
decade			Y			Y			Y
Prob > chi2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pseudo R-squared	0.075	0.094	0.138	0.102	0.118	0.153	0.105	0.121	0.153
Observations	4809	4809	4809	4809	4809	4809	4809	4809	4809

Robust standard errors in parentheses. Standard errors clustered at district level. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant.

Table 12. First-stage infant deaths regressions (large sample)

Negative binomial	(1) infant deaths	(2) infant deaths	(3) infant deaths
FP exposure 2, partial	-0.4877***	-0.4566***	-0.3278**
	(0.1528)	(0.1505)	(0.1506)
FP exposure 2, full	-0.6663***	-0.6033***	-0.1637
	(0.1555)	(0.1448)	(0.1612)
WSS pgm exposure	-0.4191***	-0.5436***	-0.3070
	(0.1433)	(0.1406)	(0.1950)
father's schooling	-0.0643***	-0.0611***	-0.0474***
	(0.0149)	(0.0136)	(0.0139)
wealth	-0.0572***	-0.0500***	-0.0580***
	(0.0119)	(0.0122)	(0.0119)
proportion of children in 1971	-0.5459	-0.1728	-0.0587
proportion enrolled in 1971	-1.4463	-0.8139	-0.9380
district		Y	Y
decade			Y
Prob > chi2	0.000	0.000	0.000
Pseudo R-squared	0.052	0.123	0.140
Observations	4809	4809	4809

Robust standard errors in parentheses. Standard errors clustered at district level.
***p<0.01, **p<0.05, *p<0.1. All regressions include a constant.

Table 13. First-stage schooling regressions (large sample)

Ordinary least squares	(1) schooling	(2) schooling	(3) schooling
FP exposure 2, partial	0.2738	0.3538*	0.1927
	(0.1808)	(0.1798)	(0.1718)
FP exposure 2, full	1.3697***	1.4945***	0.8992***
	(0.1862)	(0.1810)	(0.1823)
WSS pgm exposure	1.2456***	1.3345***	0.7715***
	(0.1694)	(0.1694)	(0.2156)
father's schooling	0.4405***	0.3818***	0.3646***
	(0.0174)	(0.0152)	(0.0151)
wealth	0.2229***	0.2164***	0.2312***
	(0.0100)	(0.0099)	(0.0100)
proportion of children in 1971	11.7575**	3.0238	2.4152
proportion enrolled in 1971	7.1075***	6.3595***	6.5655***
district		Y	Y
decade			Y
F test of excluded instruments (F statistic/P-value)	659.40/0.000	34.08/0.000	34.87/0.000
R-squared	0.490	0.542	0.553
Observations	4807	4807	4807

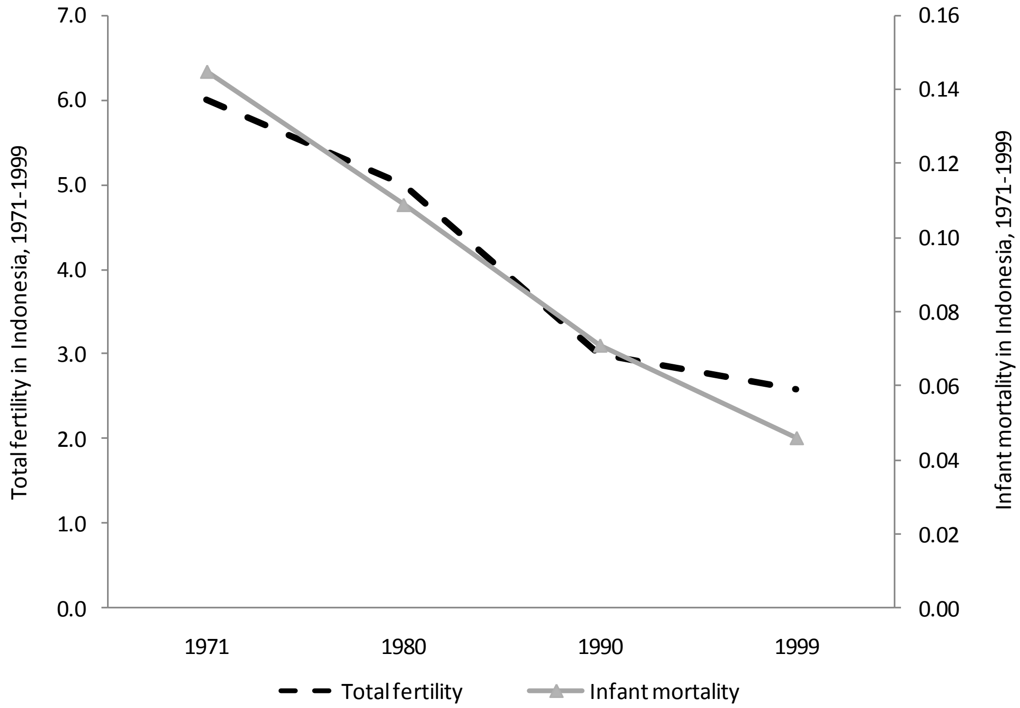
Robust standard errors in parentheses. Standard errors clustered at district level.
***p<0.01, **p<0.05, *p<0.1. All regressions include a constant.

Table 14. Second-stage total fertility regressions (large sample)

	(1)	(2)	(3)
Poisson	tfr	tfr	tfr
FP exposure 2, partial	-0.2504***	-0.2859***	-0.1941***
	(0.0531)	(0.0310)	(0.0241)
FP exposure 2, full	-0.4527***	-0.5041***	-0.2332***
	(0.0466)	(0.0354)	(0.0226)
infant deaths	0.8112**	0.3843***	0.2223***
	(0.3350)	(0.1027)	(0.0775)
schooling	-0.0450***	-0.0677***	-0.0378***
	(0.0097)	(0.0076)	(0.0057)
wealth	0.0196***	0.0224***	0.0077***
	(0.0021)	(0.0027)	(0.0023)
proportion of children in 1971	Y	Y	Y
proportion enrolled in 1971	Y	Y	Y
district		Y	Y
decade			Y
Prob > chi2	0.000	0.000	0.000
Pseudo R-squared	0.083	0.102	0.141
Observations	4809	4809	4809

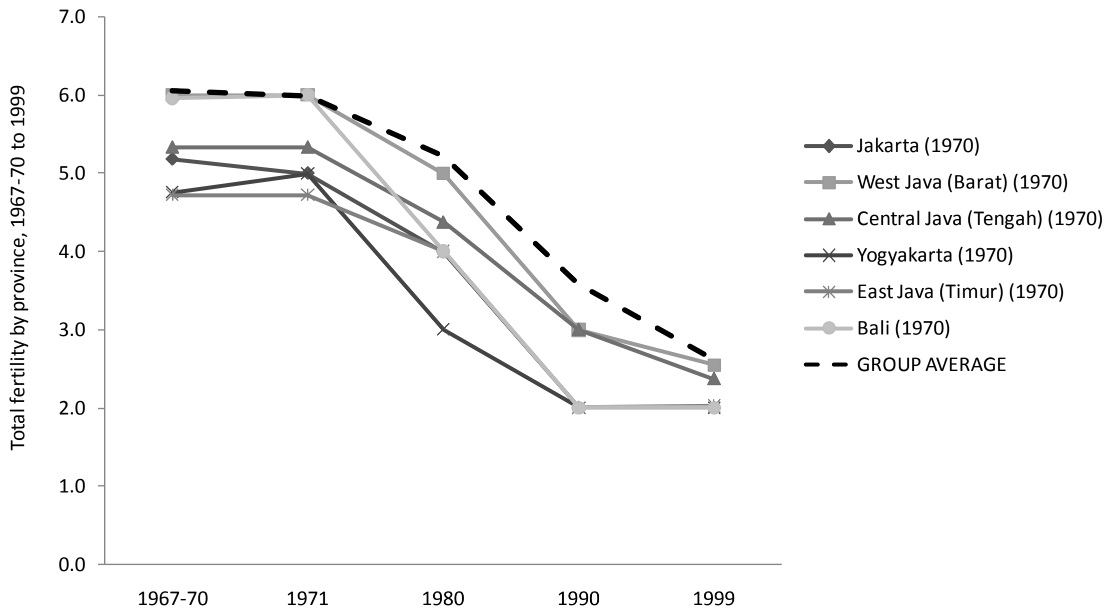
Bootstrapped standard errors in parentheses. Standard errors clustered at district level. ***p<0.01, **p<0.05, *p<0.1. All regressions include a constant. Infant deaths and schooling are instrumented.

Figure 1. Total fertility and infant mortality in Indonesia, 1971-1999



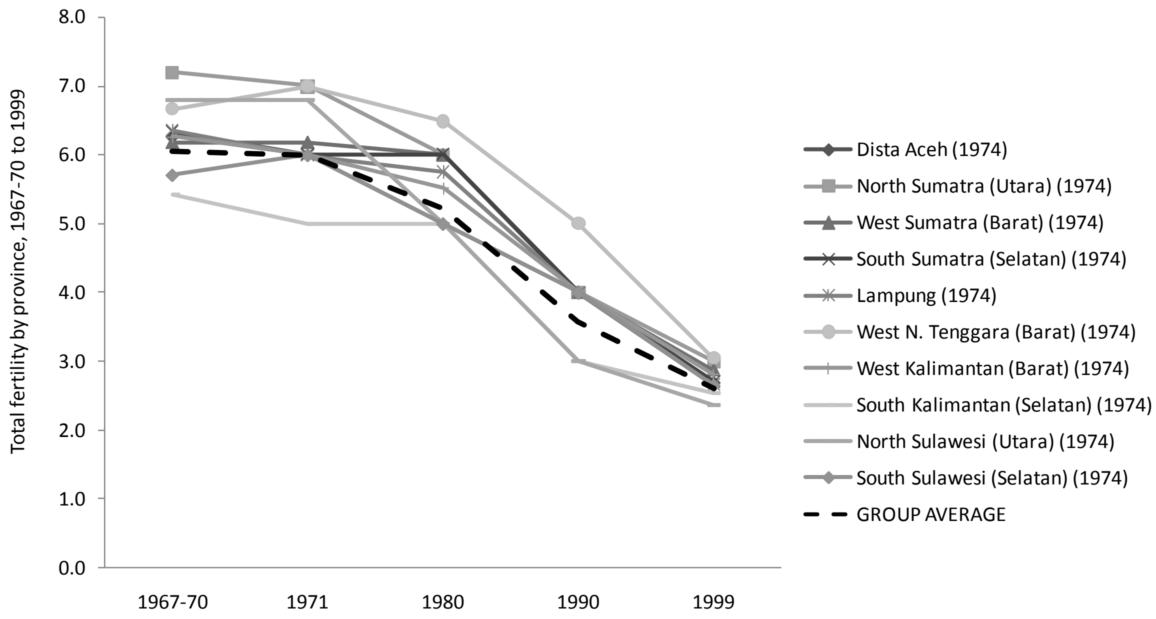
Note: Excludes Irian Jaya (Papua).
Source: (BPS 2009b, BPS 2009a).

Figure 2. Total fertility in provinces where NFPP started in 1970



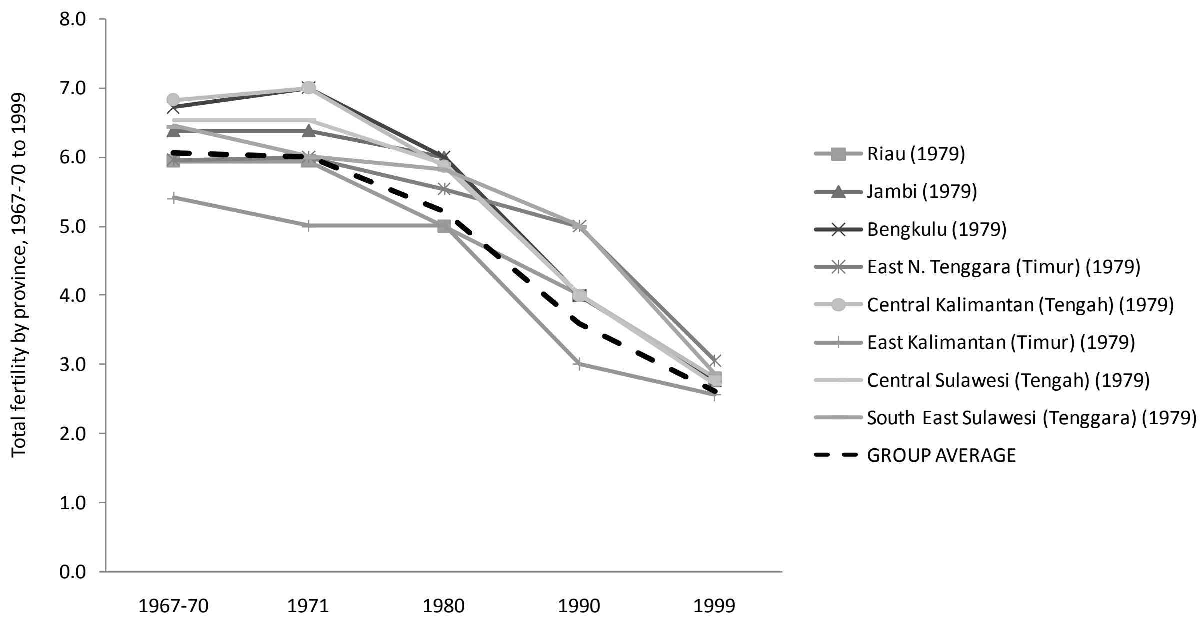
Note: Excludes Irian Jaya (Papua).
Source: (BPS 2009b).

Figure 3. Total fertility in provinces where NFPP started in 1974



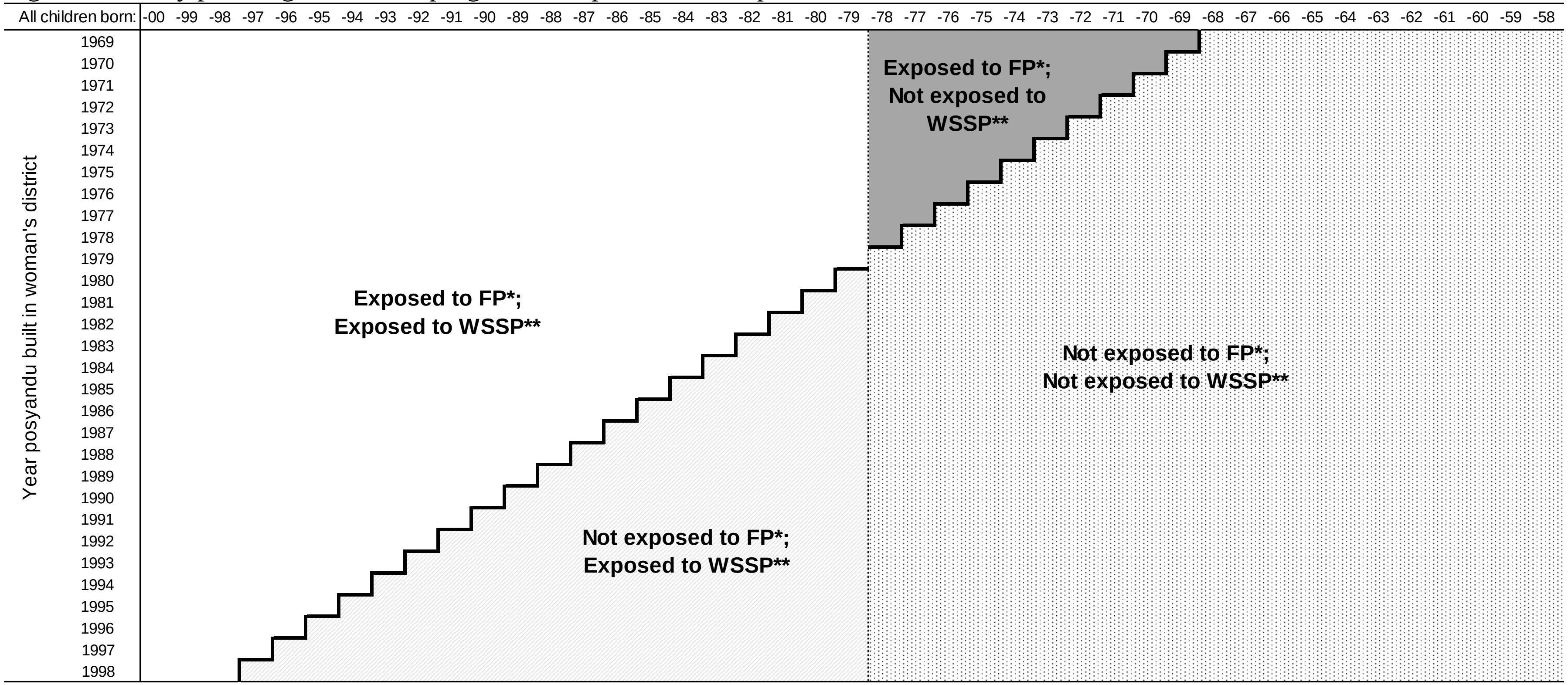
Note: Excludes Irian Jaya (Papua).
Source: (BPS 2009b).

Figure 4. Total fertility in provinces where NFPP started in 1979



Note: Excludes Irian Jaya (Papua).
Source: (BPS 2009b).

Figure 5. Family planning and WSSP programme exposure for sample women



Note: FP = family planning, WSSP = water supply and sanitation programme

*A woman is defined as partially exposed to family planning if her second and subsequent children were born after a posyandu constructed in her district (not shown). For example, if the posyandu was constructed in 1972 and a woman's first child was born in 1969, her second child in 1973 and her third child in 1975 she is defined as partially exposed to family planning.

**A woman is defined as exposed to the WSSP if all her births took place in a district where the WSSP was rolled out. In cases where not all births of a woman occurred after the start of the WSSP she is assumed to be unexposed to the WSSP.

Figure 6. Distribution of total fertility

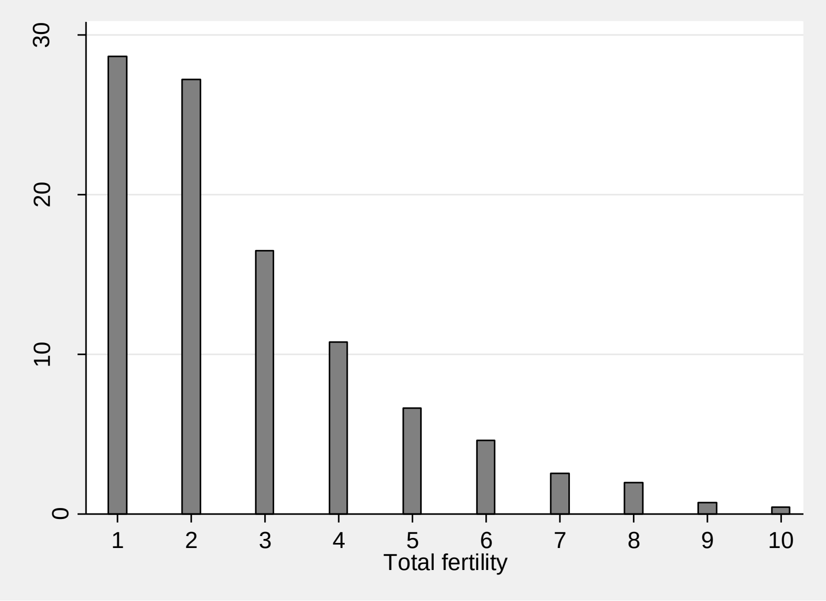
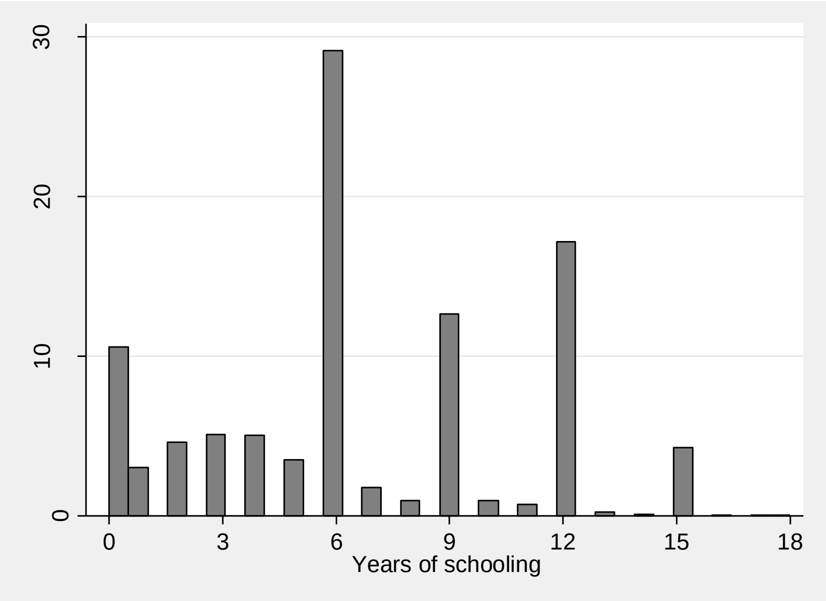


Figure 7. Distribution of schooling



Chapter 2: Does More Schooling Make You Slimmer? The Causal Effect of Schooling on BMI in Indonesia

Abstract

BMI for Indonesian men and women has been rising over the last three decades and in 2002 nearly one-in-ten men and more than one-in-five women were overweight. At the same time noncommunicable diseases, in particular cardiovascular diseases, for which overweight constitutes a major risk factor has become the main cause of death. In developing countries, schooling and BMI tend to be either positively correlated or uncorrelated but there exists no evidence on the causal effect of schooling on BMI. This chapter empirically assesses whether more schooling causes healthier BMI for men and women in Indonesia by using two instrumental variables to capture exogenous variation in schooling. The first instrument takes advantage of the massive Presidential Instruction Primary School construction programme (SD INPRES) over the period 1973/74-1978/79, the second instrument is father's schooling. Two results stand out: more schooling causes higher BMI for men and there is no causal effect of schooling on BMI for women. Such a differential effect of schooling on BMI of men and women has previously been found in developed countries. This chapter also provides some very preliminary evidence that the shift from blue collar, relatively energy-intensive occupations to relatively more sedentary white collar and service sector jobs may be one contributing factor to why more schooling increases BMI for men.

2.1 Introduction

Middle-income countries are facing a new major health problem: the rapidly rising prevalence of overweight while still having to cope with the existence of underweight (Philipson, Posner 2008, WHO 2004). Simultaneously, overweight-related, noncommunicable diseases such as cardiovascular diseases, type 2 diabetes, musculoskeletal disorders, and certain cancers, are becoming the main causes of adult deaths in middle-income countries (WHO 2004, WHO 2009, WHO expert consultation 2004).

Indonesia is no exception. Mean body mass index (BMI) both for Indonesian men and women has been rising over the last three decades (Figure 8), and nearly one-in-ten men and more than one-in-five women were overweight in 2002 (WHO 2011a). At the same time noncommunicable diseases, in particular cardiovascular diseases, for which overweight constitutes a major risk factor has become the main cause of death accounting for 24 percent of male and almost 26 percent of female deaths in 2004 (WHO 2011b).[21]

The overweight epidemic has severe adverse consequences for individuals and society. At the individual level, increased morbidity and mortality, higher private health care costs, and frequently worse employment opportunities and lower wages reduce life quality directly and indirectly while at the society level, the increasing prevalence of overweight raises public health care costs and reduces productivity (Chopra, Galbraith et al. 2004, Dor, Ferguson et al. 2010, Grossman, Kaestner 1997, Withrow, Alter 2011).

The main cause of overweight is the excess of calories consumed over calories expended due largely to increasingly high-fat, energy-rich diets and reductions in physical activity related to the sedentary nature of many jobs, changes in transportation options, and urbanization (WHO 2004, WHO 2009). The rapid growth in the prevalence of overweight-related noncommunicable diseases makes it urgent to understand which public interventions would be most effective in decelerating the overweight epidemic. The most frequently discussed public interventions are additional formal schooling; taxation to encourage consumption of healthier foods; regulation of fast food; and targeted information campaigns (Philipson, Posner 2008).

[21] All noncommunicable diseases accounted for 61 percent of deaths in Indonesia in 2002 (World Bank 2011a).

In general, more schooling is correlated with better health (Grossman, Kaestner 1997). However, evidence from studies that treat schooling as endogenous indicates that much of the positive correlation between schooling and health is due to unobserved heterogeneity rather than a causal effect of schooling (Auld, Sidhu 2005). If this is the case, policies aimed at improving health, including lowering the prevalence of overweight, through increased schooling may not be effective.

In developing countries, schooling and overweight or BMI tend to be either positively correlated or uncorrelated but there exists no evidence on the causal effect of schooling on overweight and BMI (Monteiro, Moura et al. 2004, Witloear, Strauss et al. 2011). Increasingly, evidence on the causal effect of schooling on overweight and BMI is becoming available for developed countries and where more schooling appears to have no effect on overweight or BMI except for a few cases where it is found to cause lower BMI or reduce the probability of overweight (Arendt 2005, Grabner 2009, Jürges, Reinhold et al. 2011, Webbink, Martin et al. 2010).

This chapter uses the instrumental variable approach to assess whether more schooling causes healthier BMI for men and women in Indonesia. As far as known, no previous study has assessed the causal effect of schooling on BMI in Indonesia, or in any other developing country. The first instrument takes advantage of the massive Presidential Instruction Primary School construction programme (SD INPRES) over the period 1973/74-1978/79 during which the number of primary schools in the country roughly doubled and primary enrollment rose from 13 to 19 million students. This program has previously been used but with different measures of programme exposure, to assess the impact of schooling on wages and of schooling on intergenerational educational mobility (Duflo 2001, Hertz, Jayasundera 2007). The intensity and timing of the construction of SD INPRES schools varied substantially across districts, which is used to construct the measure of SD INPRES programme exposure. The second instrument is father's schooling, which is a valid instrument if an individual's schooling is highly correlated with their father's schooling, and father's schooling only affects BMI through the schooling of their children, which it is argued is the case in this context.

To preempt the findings of this chapter two results stand out for Indonesia: more schooling causes higher BMI for men and there is no causal effect of schooling on BMI for women. This chapter also provides some very preliminary evidence that the shift from blue collar, relatively energy-intensive jobs to relatively more sedentary white collar and service sector jobs may be one contributing factor to why more schooling increases BMI for men.

The findings of this chapter adds to the emerging evidence on the causal effect of schooling on BMI in general, and as a first provides evidence on the causal effect of schooling on BMI in a developing country.

The next section provides a selective review of the literature on schooling and overweight. Section 2.3 provides the context for schooling, overweight, and overweight-related, noncommunicable diseases in Indonesia. In Section 2.4 the sample, data, and instrumental variables used for the analysis are described. Next, Section 2.5 outlines the conceptual framework and empirical methodology. Section 2.6 presents and discusses the findings. Finally, Section 2.7 concludes.

2.2 Literature Review

There is a vast literature on the relationship between schooling and health (Cutler, Lleras-Muney 2006). However, the literature on schooling and overweight, and especially, on the causal relationship from schooling to overweight is relatively new (Eide, Showalter 2011). This section provides a selective review of the empirical literature on schooling and overweight. Studies that examine the causal effect of schooling on overweight in developed countries are discussed since as mentioned above, no such studies exist for developing countries. There is also a dearth of correlational studies of schooling and overweight in developing countries (Monteiro, Moura et al. 2004, Sobal, Stunkard 1989).[22] One exception is a recent study of the relationship between schooling and BMI in Indonesia that is particularly relevant to the analysis in this chapter (Witloear, Strauss et al. 2011).

Recent studies that examine the causal effect of schooling on overweight in developed countries use various methods including instrumental variables, regression discontinuity designs, matching, and twin data. Most studies use changes in compulsory schooling laws or in state-level education policy to capture exogenous variation in schooling.

Apart from two studies (Conti, Heckman 2010, Grabner 2009), the existing literature does not find that more schooling causes healthier BMI in developed countries

[22] Some sub-national level correlational studies that examine cities or districts within a developing country exist ((Monteiro, Moura et al. 2004)).

(Arendt 2005). Most of the studies that use changes in compulsory schooling laws or state-level education policy as instruments find that the instrumental variable estimates are a great deal larger than those obtained by ordinary least squares. A plausible explanation is that the return to schooling is larger for those who changed their schooling behaviour in response to the change in schooling laws or education policy than for individuals who did not (Eide, Showalter 2011).

In the United States, changes in compulsory state-level schooling laws imposed a minimum amount of schooling required to apply for a work permit, which forced some individuals to remain longer in school, about 2-9 months, than they otherwise would have and these changes in schooling laws is used as an instrument for schooling (Grabner 2009). The samples consist of approximately 11,900 men and women who are 18 years or older. In the first-stage regressions, the F-test for the excluded instruments being jointly zero ranges from 9 to 14 indicating that the instruments are relevant. According to the estimates, additional schooling causes a reduction in BMI and in the probability of being obese both for men and women, although the effect is larger for women. Comparing the ordinary least squares estimates to the instrumental variable estimates the latter are about three times as large.

Another study for the United States uses changes in state-level education policy as instruments for schooling but find no significant impact of schooling on overweight nor obesity for either men or women (Kenkel, Lillard 2006). The samples consist of 3,248 adult males and 3,275 adult females. The instruments are based on the education environment individuals faced at the time of their schooling and capture the difficulty of graduating from high school; the difficulty of obtaining General Educational Development High School Equivalency Diploma (GED) certification; and per capita education spending. Parental schooling is also included as in instrument in some of the models. The instruments are highly relevant with the first-stage F-test of relevance ranging from 21 to 64. Based on the ordinary least squares estimates treating schooling as exogenous, men who graduate from high school are significantly more likely to be overweight and so are men who received GED certification. By contrast, for women there is no significant effect of either. The instrumental variable estimates are not significant for men nor for women, which implies that more schooling does not cause healthier BMI.

In 1958, Danish education reforms removed barriers to entry to higher education for children from disadvantaged backgrounds and living in underserved areas, and in 1975 the school-leaving age was raised, which enables the construction of instruments that capture exogenous variation in schooling (Arendt 2005). The effect of schooling on overweight and on obesity is examined separately for 3,420 men and for 3,096 women. The increase in mean schooling attainment was larger for the 1958 than the 1975 education reforms and in the first-stage schooling regressions only the dummy variable for the 1958 reform is significant. Moreover, the F-statistic for the excluded instruments being jointly zero is only 4 for women and 5 for men, suggesting that the instruments are weak with implications for bias of the estimates (Staiger, Stock 1997). More schooling is positively but never significantly related to healthier BMI for men or women, and the ordinary least squares estimates are smaller in size than the instrumental variable estimates.

A study of adult men and women in 1999 and 2003 uses the increase in the number of academic-track (grammar) schools in Germany, which varied across states and over time, as an instrumental variable to estimate the causal effect of schooling on obesity (Jürges, Reinhold et al. 2011). Since academic track schools generally comprise nine years compared to only six years of schooling for the other two types of secondary schools in the country, the increase in the number of academic track schools led to additional years of schooling for a substantial share of the population. According to the ordinary least squares estimates, men and women with more schooling have a significantly lower probability of being overweight and of being obese. In the first-stage regressions, the instrument is found to be relevant for both men and women with F-statistics between 13 and 30. In the second-stage, the instrumental variable estimates, suggest that more schooling significantly increases the probability of being overweight and of being obese for men, and the effect is relatively large, but it is not significant for women. It is argued that this result is due to lifestyle changes conducive to overweight as more schooled men move from relatively physical blue collar jobs to relatively sedentary white collar jobs.

Another study for Germany uses the same data and age group but a different instrument: the abolition of secondary school fees, to assess the causal effect of schooling on overweight and on obesity (Reinhold, Jürges 2010). Fees were abolished across different states at different times, which is used to capture exogenous variation in schooling. Similar to the other German study discussed above (Jürges, Reinhold et al. 2011), the ordinary least squares estimates indicate that more schooling is significantly associated with a lower probability of being overweight and of being obese for both men and women. In the first-stage schooling regressions the F-statistics for instrument relevance range from 13 to 15 indicating that the instrument is relevant. But based on the instrumental variable estimates schooling does not significantly affect the probability of being overweight or of obese, and again the schooling coefficient switches from negative for the ordinary least squares estimates to positive for the instrumental variable estimates.

Changes in compulsory schooling laws regarding school-leaving age in the United Kingdom in 1947 and 1972 led to large differences in schooling attainment for individuals born only months apart, which are used to assess the causal effect of schooling on various health measures including overweight (Clark, Royer 2010). The 1947 change in schooling laws meant that roughly half of the cohorts affected by the change obtained one additional year of schooling and for the 1972 change approximately a quarter of the affected cohorts gained one additional year of schooling. These changes in schooling attainment allow the use of regression discontinuity methods to assess the causal effect of schooling on overweight and on obesity. The sample consists of roughly 16,500 female and male adults born in the 15-year interval around each of the 1947 and the 1972 school law changes. The ordinary least squares estimates suggests that more schooling lowers the probability of being overweight and of being obese although the effects are relatively small. By contrast, according to the instrumental variables estimates schooling does not affect the probabilities of being overweight or of being obese, and the coefficients sometimes switch sign from negative to positive.

A recent study of schooling and overweight uses a model that allows for sorting based on cognitive and non-cognitive ability and health endowments into different schooling levels to examine the impact of schooling on obesity (Conti, Heckman 2010). The effect of schooling on obesity is assessed for 3,777 men and 3,620 women who were 30 years old in 2000. The main result is that schooling is much more important in explaining the obesity gap for men than for women for whom the obesity differences by schooling are found to be entirely the result of selection.

One non-causal study is reviewed as it is highly relevant because it exmines the relationship between schooling and overweight in Indonesia (Witloear, Strauss et al. 2011). The study documents, among other things, changes in the weight distribution in Indonesia for men and women age 45 years and more for the period 1993-2007 using the four rounds of the Indonesia Family Life Survey (IFLS). The male sample consists of 12,837 observations and the female sample of 14,735 observations. The proportion of overweight in the population aged 45 years and above has risen rapidly from 8.5 percent of men in 1993 to 17 percent in 2007, and for women from 14 percent to 31 percent. To examine the relationship between schooling and overweight the four rounds of the IFLS are pooled and a set of multivariate regressions are estimated using ordinary least squares. Schooling is captured by dummy variables for some primary schooling, complete primary schooling or more, and complete junior secondary school or more. Men who have completed junior secondary school have significantly higher BMI than men with no schooling. Among women, those with some primary schooling have significantly higher BMI than those with no schooling, and women with complete primary schooling or more have significantly higher BMI than women with some primary schooling. However, there is no additional effect of having completed junior secondary school or more for women.

These findings are in contrast to those for developed countries that typically find that more schooling is associated with lower BMI when treating schooling as exogenous. Given that there are no studies for developing countries including Indonesia on the causal effect of schooling on overweight a knowledge gap that this chapter is an attempt to address.

2.3 Indonesia Context: Schooling and Health

In the early 1970s before the massive primary SD INPRES school construction programme began, Indonesia had a population of 104 million living in 26 provinces on approximately 3,000 islands speaking roughly 250 different languages (Beeby 1979).

2.3.1 Primary Schooling System

The Indonesian school system consisted of six years of primary school; three years of junior secondary school; and three years of senior secondary school. The legal minimum school starting age was six, but the majority of children began school at age seven and some at an even older age, primarily due to shortages of school places. The typical primary school had one principal and five to six teachers and between 150-300 students (Beeby 1979).

Indonesia had a relatively high capacity to enroll primary school aged children in 1971, with a gross primary enrollment ratio of 80 percent. However, due to large numbers of over-age students, only 68 percent of eligible children (the net primary enrollment ratio) were enrolled and rural children and girls were much less likely to be enrolled than urban students and boys respectively (World Bank 1975). Moreover, dropout rates were high. 13 percent of children enrolled in primary school dropped out before grade three; 24 percent before grade four; and more than 55 percent before grade six. Out of those reaching grade six, only 61 percent graduated (Beeby 1979).

The majority of parents interviewed in education surveys cited lack of money as the main reason for their child dropping out. After the cost of schooling, lack of awareness of schooling opportunities was the most commonly cited reason for dropping out. The cost of sending a child to school including school fees imposed a barrier to entry for children from poor households, and when parents sent their children to school, they were frequently unable to keep them in school (Daroesman 1972).

There were two types of primary school fees in the early 1970s, entrance fees and monthly fees collected by parents' association for public schools. To increase transparency and accountability and reduce fees for the poor, a fixed, single parental contribution (SPP) was introduced in 1971 to replace the old fee system (Beeby 1979). The SPP comprised 1-2 percent of household income depending on costs in the area in which the household resided (World Bank 1975). In addition, there were other direct costs for uniforms, school supplies, and transport, and the opportunity cost of sending children to school. Primary school fees were abolished for grades 1-3 in 1977 and for grades 4-6 in 1978 (Chernichovsky, Meesook 1985). As well as the high cost of schooling and lack of awareness of schooling opportunities, many parents considered three to four years of schooling adequate, feeling that the literacy and numeracy skills achieved at these schooling levels were sufficient for the needs of their children (Beeby 1979).

The constraints of the primary school system were evident in schooling outputs and outcomes. In 1970, average years of schooling of the adult population (15 years and older) was only 2.8 years; 41 percent of the population had no schooling at all; and just 19 percent had completed primary schooling (World Bank 1975, Barro, Lee 2010). Only 60 percent of the adult population was literate. There were also large differences by location and gender: 79 percent of the urban population was literate but only 55 percent of the rural population and merely 49 percent of women compared to 71 percent of men (Table 15).

2.3.2 Overweight Prevalence and Chronic Diseases

A normal BMI lies in the range 18.5-24.99; a BMI below 18.5 indicates underweight; whereas a BMI in the range 25.0-29.99 indicates overweight (pre-obese) and a BMI of 30.0 and above obesity (WHO 2011a). These are general cut-off points that may vary across different populations. In particular, "Asians generally tend to have a higher percentage of body fat than white people of the same age, sex, and BMI. Also, the proportion of Asian people with risk factors for type 2 diabetes and cardiovascular disease is substantial even below the existing WHO BMI cut-off point of 25 kg/m^2" (WHO expert consultation 2004: 161). Some evidence suggests that it would be appropriate to lower the existing BMI cut-off points for Indonesians by three units (WHO expert consultation 2004).

BMI has risen since the 1980s in Indonesia because of changes in diet and physical activity. Figure 8 shows that mean BMI for men and women has risen steadily over the last three decades, and somewhat faster (starting from a higher level) for women. Based on the standard BMI cut-off points, in early 2000, 17.3 percent of Indonesian women were overweight and 3.6 percent obese (Figure 9). Among Indonesian men, 8.4 percent were overweight and 1.1 percent obese (WHO 2011a). At the same time, 15 percent of the Indonesian population was still underweight (World Bank 2011a). If the standard BMI cut-off points were adjusted down as proposed by emerging evidence, even larger proportions of the Indonesian population would be categorized as overweight and obese respectively.

To put the prevalence of overweight in perspective data are also shown for China and India, two other large middle-income countries (Figure 9). In China, relatively more men and women are overweight (19.1 percent and 18.1 percent respectively) and relatively more men obese (2.4 percent) while the share of obese women is somewhat smaller (3.4 percent) than in Indonesia (WHO 2011a). In India, the prevalence of overweight and of obesity both for men and for women is lower than in Indonesia.

In low- and middle-income countries, including Indonesia, six risk factors account for nearly one in five deaths: high blood pressure, high blood glucose, physical inactivity, overweight and obesity, high cholesterol, and low intake of fruit and vegetables and most of these factors are related to overweight (WHO 2009). These risk factors are responsible for cardiovascular diseases, stroke, and type 2 diabetes that are three of the main causes of death in Indonesia. In 2002, noncommunicable diseases accounted for 61 percent of all deaths in the country (WHO 2011b), and among these, cardiovascular diseases alone accounted for 24 percent of male deaths and 26 percent of female deaths, and type 2 diabetes for 2 percent of male deaths and 3 percent of female deaths (Figure 10).[23]

The fact that noncommunicable diseases are the most important cause of death, the majority of which are directly or indirectly related to overweight, combined with the rising prevalence of overweight in Indonesia makes it vital to understand what factors influence BMI, and in particular, if more schooling causes a healthier BMI.

2.4 Sample Characteristics and Data

This section describes the sample, data, and instrumental variables used to assess the causal effect of schooling on BMI for men and women in Indonesia. Variable descriptions and summary statistics for all variables are provided in Table 17, Table 18, and Table 19.

2.4.1 Sample

The data used for the analysis come from the Indonesia Family Life Survey (IFLS), a repeated socioeconomic and health survey run by the RAND Corporation and the Center for Population and Policy Studies of the University of Gadjah Mada in Indonesia. The first three rounds of the survey: IFLS1, IFLS2, and IFLS3 conducted in 1993, 1997 and 2000 respectively are used. The 10,435 households interviewed in IFLS3 live in 13 of Indonesia's 26 provinces, and represent approximately 83 percent of the Indonesian population in 1993 (Strauss, Beegle et al. 2004). The 13 provinces covered by the IFLS are: Sumatra Utara, Sumatra Barat, Sumatra Selantan, Lampung, DKI Jakarta, Java Barat, Java Tengah, DI Yogyakarta, Java Timur, Bali, Nusa Tenggara Barat, Kalimantan Selantan, and Sulawesi Selantan, which are the most densely populated areas in Indonesia (Annex figure 1**Error! Reference source not found.**).

The effect of schooling on BMI is examined separately for men and women to account for the important role of gender in explaining differences in health (Conti, Heckman 2010). The unit of analysis is men and women aged 20-50 years in 2000. This age group was chosen to include individuals not exposed and exposed to the SD INPRES programme that started in 1973/74, which is used as an instrumental variable,

[23] Cardiovascular diseases include: rheumatic heart disease, hypertensive heart disease, ischaemic heart disease, cerebrovascular disease, and inflammatory heart disease (WHO 2011b).

and that were adults at the time of the survey. 8,303 men and 8,775 women were 20-50 years old at the time of the survey in 2000, this is reduced to 4,300 men and 4,821 women after removing observations for which all variables required for the analysis are not available. Mean schooling of fathers and SD INPRES programme exposure are significantly higher for men and women in the samples used for the analysis than for the excluded observations. Therefore, results may not be valid outside the sample.

The data on SD INPRES programme school construction, the INPRES water supply and sanitation programme, and the primary school enrolment rate in each district in 1971 come from administrative records for 1973-1979 (BAPPENAS 1973, 1974, 1975, 1976, 1978). The IFLS data were merged with the administrative data, which was very time consuming as several of the 283 districts (kabupaten) changed name, some more than once, between 1973/74 when the SD INPRES programme started and 2000 when the third IFLS wave was conducted.

2.4.2 Body Mass Index

The health variable of interest is body mass index defined as weight in kilograms divided by height (in metres) squared. BMI is not self-reported but measured by the survey enumerators, which reduces the scope for measurement error. In the sample, based on the standard BMI cut-off points (Section 2.3), women are more prone to be overweight than men, and the proportion of overweight women is larger than the proportion of underweight women. Among women, 21 percent are overweight, and within this group, 3.6 percent are obese (Table 21). By contrast, a smaller share, 11 percent of men are overweight, and among these, 1.6 percent is obese. When it comes to underweight, nearly 14 percent of women in the sample are underweight compared to 16 percent for men. Figure 12 shows the empirical cumulative distributions of BMI for the female and male samples, which further illustrate that relatively fewer men than women are overweight and the high prevalence of overweight in the sample. Among the women in the sample mean BMI is 22.2 with a standard deviation of 3.7 compared to a mean BMI of 21.2 for men with a 3.1 standard deviation (Table 18).

2.4.3 Schooling

Based on the existing evidence, the effect of schooling on BMI is ambiguous and appears to differ by gender. In this chapter, schooling is defined as years of formal schooling and is constructed by combining the reported highest grade and level of formal schooling completed. Women on average have less schooling than men. Mean years of schooling for women is 7.5 with a standard deviation of 4.3 years compared to 8.5 years with a standard deviation of 4.0 years for men. Figure 13 shows the heaping of schooling around the completion of each schooling level. Almost 9 percent of women and 4 percent of men have no schooling; 27 percent of women and 24 percent of men have complete primary schooling; 13 percent of women and 14 percent of men have completed junior secondary school; 22 percent of women and 27 percent of men have completed senior secondary school; and 6 percent of women and 7 percent of men have a university education.

Examining schooling for men and women exposed and not exposed to the SD INPRES programme reveals differences across the two groups. For exposed men mean schooling is higher: 9.3 years relative to 6.8 years for men not exposed to the SD INPRES programme. Similarly, mean schooling is higher for exposed women: 8.7 years compared to 5.1 years for unexposed women (Table 20). The differences in schooling across the exposed and unexposed groups are significantly different, and indicate that the SD INPRES programme, unconditional on any other factors that may influence schooling, is associated with more schooling in Indonesia.

Schooling is endogenous in the analysis of health (Grossman 2006), including BMI. For instance, unobserved heterogeneity such as innate ability or preferences may influence schooling and BMI in the same direction (Grossman, Kaestner 1997, Behrman 1997, Sander 1992). Causation between schooling and health outcomes may run in both directions, which would also bias estimates of the effect of schooling on health (Grossman, Kaestner 1997). This may even be the case for adults who have completed their schooling if past health is not directly controlled for (Grossman 2006). Schooling may also be measured with error as recall periods become longer and lead to biased estimates (Griliches 1977, McCrary, Royer 2011).

Different approaches can be used to address the endogeneity of schooling including controlling for past health although such data are hard to come by; by using twin or sibling data although these samples are usually small, or by finding valid instrumental variables, which is often difficult (Grossman 2006). To address the endogeneity of schooling this chapter uses SD INPRES programme exposure and years of schooling of fathers as instruments (see below).

2.4.4 Household Wealth and Location and Marital Status

Being wealthier is generally associated with higher BMI in developing countries, primarily because wealthier individuals can afford a higher food intake and arguably because of social norms that consider overweight a sign of health, affluence, and prestige (Sobal, Stunkard 1989, Graham, Felton 2005).[24] However, overweight appears to be becoming more common also among poorer populations in developing countries over time, especially, in urban areas (Monteiro, Moura et al. 2004)

In this chapter, household wealth is captured by three dummy variables indicating the poorest, middle, and richest tertiles of households based on an asset index. The index is constructed based on access to or ownership of six items: electricity; piped water; toilet with a septic tank; refrigerator; electric or gas stove; and television set. Each item is assigned a weight equal to the inverse of the fraction of households that own it so that rarer items receive a larger weight. Household wealth is used rather than household income or consumption since asset holdings tend to change less quickly over time. Moreover, it helps minimize recall bias and measurement error commonly associated with income and consumption (McKenzie 2005). The mean asset index for the female sample is 6.1 and for the male sample 6.0 (Table 18). Among the poorest tertile of women, the mean asset index is 1.6 and among men 1.7; for the middle tertile of women 4.8 compared to 4.5 for men, and for the richest tertile of women 13.7 relative to 12.5 for men. Mean BMI increases with wealth tertile both for men and for women suggesting that without controlling for any other relevant factors, higher wealth is associated with higher BMI.

Whether an individual lives in a rural or urban area influences BMI, predominantly through differences in diet and physical activity. Urban residents generally have lifestyles that are more sedentary and diets that are higher in fat and energy, both of which contribute to overweight (WHO 1998). Among the men and women in the two samples, 48-49 percent lives in urban areas.[25]

The effect of marital status on BMI may go in either direction. On the one hand, individuals with lower BMI may be more likely to select into marriage, or married individuals be more likely to have a healthy BMI due to social support provided by their partner. On the other hand, those who are married may have a higher BMI if their social

[24] The opposite relationship generally holds in developed countries (Sobal, Stunkard 1989).
[25] For the country as a whole, 58 percent of the population lives in urban areas and 42 percent in urban areas (World Bank 2011b).

environment encourages eating more often and exercising less, or if they face less of an incentive to maintain a healthy BMI as they are no longer searching for a partner (Averett, Sikora et al. 2008).[26] In the sample, the majority of men and women are married, and relatively more women (77 percent) than men (68 percent) (Table 18), which is as expected because women tend to marry at a younger age than men in Indonesia (United Nations 2010).

2.4.5 District-Specific and Time Effects

To account for district-specific effects that may affect BMI the model includes dummy variables for each district. Likewise, to control for year-specific effects, including the observed positive trend in BMI, dummies for year of birth are used. Interactions between year of birth and district primary enrollment rates in 1971 are included in the model to account for targeted SD INPRES programme allocation (Section 2.5). Moreover, interactions of year of birth and the rollout of the public water supply and sanitation programme across districts simultaneous to the SD INPRES programme are used to account for any direct effects of this programme on BMI.

2.4.6 Instrumental Variables

The SD INPRES Programme

The number of primary schools in Indonesia increased gradually over the period 1956-1965, but the construction effort slowed down after 1966 (UNESCO 1976). Indonesia's first national five-year development plan 1968/69-1973/74 (Repelita I) focused on economic development and its education component was concerned almost exclusively with vocational and technical training, while ignoring the overall structure of education including primary education (Beeby 1979, World Bank 1989).

In 1971, two years before the start of the SD INPRES programme, there were approximately 61,000 public primary schools under the jurisdiction of the Department of Education in the country, and in addition, there were about 20,000 Madrasah primary schools (World Bank 1975).[27] Out of school-age children, 68 percent were enrolled, and the number of children nearing school age was increasing rapidly (Beeby 1979).

The hike in oil prices in 1973/74 expanded the size of the Indonesian government

[26] Emerging evidence for developed countries increasingly points to a positive relationship between marriage and BMI (Sobal, Rauschenbach et al. 2003).
[27] Most of these schools had a status similar to that of the schools under the Department of Education and followed the state curriculum.

budget 2.5 times between 1973 and 1975, which substantially increased the Government of Indonesia's (GOI) ability to finance the extensive education reforms that were part of the second national five-year development plan (Repelita II) (World Bank 1989).[28]

Table 16 shows the number of primary schools to be constructed as part of the SD INPRES programme: 62,000 schools or equivalently 2 schools per 1,000 children aged 5-14 in 1971. This constituted more than a doubling of the existing number of primary schools under the jurisdiction of the Department of Education in just five years.[29] The expansion of the school supply was staggered with the number of schools constructed gradually increasing over the course of the SD INPRES programme.

Each of the new schools had three classrooms, furniture, equipment, and toilets and textbooks were provided (UNESCO 1976). To ensure that the quality of schooling did not suffer due to teacher shortages as primary enrollment increased the GOI recruited and trained a large number of teachers. At the time, 25,000 new teachers entered the labor market annually and there were thousands of teachers not working as teachers due to the hiring freeze imposed in 1968 (World Bank 1989). The freeze was abolished in 1973, the year the SD INPRES programme started, allowing the hiring of 18,000 new teachers and 41,000 temporary and part-time positions to be made permanent. As a result, student-teacher ratios remained nearly constant between 1973/74 (31.5:1) and 1978/79 (31.8:1) (Beeby 1979, World Bank 1989). Combined with the provision of textbooks this indicates that the quality of primary schooling did not deteriorate during the SD INPRES programme (World Bank 1989, Duflo 2001).

From 1960 until the start of the SD INPRES programme in 1973/74, average annual growth in primary enrollment was only 0.9 percent per year (World Bank 1989). However, during course of the SD INPRES programme the number of students enrolled in primary school increased dramatically from 13 million in 1972/73 to 19.1 million in 1978/79 (Figure 11).

Primary school fees were abolished for grades 1-3 in 1977 and for grades 4-6 in 1978 (Chernichovsky, Meesook 1985). The removal of official school fees may also have contributed to the rise in enrollment after 1977/78 but as Figure 11 shows, the rise in primary enrollment was steady before and for about one year after fee abolition when

---

[28] Under Repelita II Rp436 billion were allocated to education compared to Rp36.6 billion under Repelita I (World Bank 1989).

[29] The number of schools planned corresponded very closely to the number of schools constructed (Beeby 1979, World Bank 1989).

it slowed down.

The intensity and timing of the construction of SD INPRES schools varied substantially across districts. An individual's programme exposure is determined by the intensity of school construction in her district of schooling and age at the time SD INPRES schools were constructed in her district (Duflo 2001).

The SD INPRES programme has been used to evalute the effect of increased school supply on schooling and wages for men (Duflo 2001). An interaction term between a dummy variable indicating age in 1974 (year first SD INPRES schools completed) and programme intensity in district of birth (number of SD INPRES school built per 1000 children aged 5-14 in 1971) is used to capture SD INPRES programme exposure based on the 1995 intercensal survey of Indonesia (SUPAS).[30] Using this variable a district in which all SD INPRES schools were constructed in 1974, for example, is assigned the same programme intensity as a district in which same number of schools were constructed, but not until 1978 (Hertz, Jayasundera 2007).

The effect of the SD INPRES programme on intergenerational education mobility has also been examined using IFLS data (Hertz, Jayasundera 2007). The programme exposure variable in this case is the average number of SD INPRES schools constructed before each of the six years of primary schooling for an individual, assuming that she attended primary school between ages 7 and 12. This variable takes advantage of the fact that schools were constructed at different times in different districts thereby capturing additional variation in SD INPRES exposure.

The variation in SD INPRES school construction intensity by district and over time also informs the measure of programme exposure used in this chapter. Because the number of SD INPRES schools constructed by district was an increasing function of time the measure of programme intensity is the cumulative number of SD INPRES schools constructed per 1,000 children aged 5-14 years in a given district in the year before an individual started primary school. This programme intensity measure uses the fact that individuals born later in a given district, conditional on them being of appropriate age when the schools were constructed, have higher SD INPRES programme exposure.

[30] District of birth rather than district of schooling is used due to data limitations. However, it is district of schooling (highly correlated with district of birth) that determines programme exposure together with year of birth.

The likely endogeneity of schooling was discussed above. To address it, SD INPRES programme exposure is used together with father's schooling (see below), to capture exogenous variation in schooling. For the SD INPRES programme to be a valid instrument it must be highly correlated with the endogenous variable schooling (relevance) and it must not be correlated with the error term in the BMI equation (exogeneity). Given the massive size of the SD INPRES programme and the rise in primary enrollment simultaneous with its rollout, it is feasible to assume that the programme affected enrollment and thereby schooling and therefore is a relevant instrument. For the SD INPRES programme to be exogenous it must only affect BMI through schooling, this seems a plausible assumption since the programme was allocated at the district level and as it would be difficult to think of channels other than schooling through which school construction may affect adult BMI. The issue of non-random programme allocation is discussed below. The relevance and validity of the instrumental variables are formally tested in Section 2.6.

An advantage of the IFLS data is that individuals who moved from their district of birth between birth and age 12 can be identified. In the female and male samples all individuals still lived in their district of birth at age 12, the age at which they are expected to complete primary school, which is important to correctly assign SD INPRES programme exposure.

The issue of children who were over-age, i.e., older than seven when starting school, needs to be considered when constructing the SD INPRES programme exposure variable. The IFLS contain data on the age at which an individual started school, which could potentially be used to determine programme exposure for over-age individuals. However, among over-age school-starters there is substantial variation in the years of schooling completed, with a large proportion not completing the full six years of primary school, and a large fraction only obtaining three years of schooling or less. This makes it difficult to assign over-age school-starters correct programme exposure. Therefore, individuals who were 8-12 years old in the year in which SD INPRES schools were constructed in their district of schooling are excluded. Thus, individuals unexposed to the SD INPRES programme are those 13 years or older in the year SD INPRES schools were built in their district, and exposed individuals are those seven years or younger at the time of school construction.

As a robustness check, the regressions are also run for a sample that includes children who were eight years old in the year SD INPRES schools were constructed in their district and treats them as fully exposed to the programme (Section 2.6).

An important aspect of the SD INPRES programme was that the new primary schools were to supplement rather than replace existing primary schools. The main objective was to provide school places for children who previously had no school to attend (Beeby 1979, Aziz 1990). Thus, the number of new schools to be built in each district was proportional to the number of children of appropriate school age not enrolled in 1972 (Chernichovsky, Meesook 1985, Aziz 1990). The location of schools was chosen by the district chief executive (bupati), after consultation with the heads of sub-districts and primary school inspectors based on three criteria. Schools were to be located where children were currently unable to start first grade in existing schools; preference should be given to urban areas where the population was mostly low-income; and preference should be given to remote areas (UNESCO 1976). Because the main determinant of the allocation of SD INPRES schools to each district was the number of school-age children not enrolled in primary school before the programme started, the estimates may be biased due to correlation between prior enrollment rates and the programme. To control for the school allocation rule an interaction term between year of birth and the proportion of children enrolled in 1971 in each district is included in the model.

The average number of SD INPRES schools constructed across all district was 2.2 per 1,000 children over the period 1973/74-1978/79.[31] For the men in the sample, mean SD INPRES programme exposure is 0.92 and for women 0.93 (the maximum cumulative number of SD INPRES schools constructed per 1,000 children in any district is 6.4) (Table 18).[32] Table 20 presents mean years of schooling, father's schooling, and BMI broken down by individuals exposed and not exposed to the SD INPRES programme by gender. Both for exposed men and women mean programme exposure is 1.4 schools per 1,000 children. Women exposed to the SD INPRES programme on average, have 3.6 years more schooling and their mean BMI is 1.6 points lower and their fathers have 2 years more of schooling than women not exposed. Similarly, exposed men have on average, more schooling (2.4 years), lower BMI (1.2

[31] For the exposed group, the minimum number of SD INPRES schools constructed per 1,000 children for a district was 0.6 and the maximum 7.7.

[32] By construction, programme exposure of the unexposed group is zero. The reasons for the difference in the mean number of schools built per 1,000 children across districts and mean SD INPRES programme exposure is that the latter includes unexposed indviduals and uses the cumulative number of schools constructed in the year before starting school per 1000 children as the numerator and the number of children in that year as the denominator whereas the former uses the period total (1973/74-1978/79) number of schools as the numerator and the number of children in 1971 as the denominator.

points), and their fathers have more schooling (2 years). All differences are statistically significant for both men and women.[33]

Father's Schooling

Father's schooling is defined as formal years of schooling completed and was obtained by linking individuals in each IFLS wave with their parents. Mean years of schooling of fathers are similar for the female and male samples: about five years with a standard deviation of 3.8-3.9 years (Table 18). In the female sample, 23 percent of fathers have no schooling; 35 percent have completed primary schooling; and close to 9 percent have completed junior or senior secondary schooling respectively. The picture is roughly comparable for the male sample: 20 percent of fathers have no schooling; 35 percent have completed primary schooling; 8.5 percent junior secondary schooling; and 8 percent senior secondary schooling.

Father's schooling is used as an instrument for schooling.[34] This instrument is valid if an individual's schooling is highly correlated with their father's schooling, and father's schooling only affects BMI through the schooling of their children. Father's schooling is not a valid instrument if it is correlated with ability and the latter is not appropriately controlled for (Card 1999). In this case, the IV estimates would have an upward bias. However, even if father's ability is correlated with his schooling and in turn correlated with his children's ability, any bias stemming from this channel is likely to be relatively minor. This is because during the period when fathers in the sample were of school-age, economic factors and social class would arguably have been more important in determining schooling than ability as this was before the start of the SD INPRES programme that extended primary schooling opportunities to the general public.

Another challenge to the validity of father's schooling as an instrument is that parents with more schooling may invest more in their children's health and health knowledge (Kenkel, Lillard 2006). In the case of schooling and BMI, arguably most of the effect of fathers' schooling works through their children's schooling. In particular, during the time period in which the individuals in the sample were brought up

[33] The fathers of those in the exposed groups tend to have more schooling because mean schooling has risen over time, and individuals in the unexposed group and therefore their parents are older than those in the exposed group.

[34] Father's schooling (together with occupation and mother's schooling) has previously been used as an instrument for schooling to estimate the causal effect of schooling on blood pressure in the United States (Berger, Leigh 1989).

overweight was uncommon in Indonesia making it highly unlikely that any health knowledge imparted at home would have been concerned with how to refrain from gaining weight regardless of father's (or mother's) schooling levels. The validity of father's schooling as an instrument is formally tested in Section 2.6.

2.5 Conceptual Framework and Empirical Methodology

This section begins by outlining the conceptual framework that underlies the empirical model. It then discusses the estimation methods and the first-stage schooling and second-stage BMI equations to be estimated.

2.5.1 Conceptual Framework

The demand for health framework on which the empirical model is based distinguishes between commodities defined as "fundamental objects of choice" and regular market goods (Grossman 1972: 224). In this framework, individuals produce commodities using a combination of their time and market goods and services. Individuals are endowed with an initial stock of health that depreciates over time and that can be improved through investment.

Each individual has household production functions for each commodity where the inputs are time and market goods and the efficiency in producing health outputs depend on individual characteristics. For instance, an individual may choose how much time to allocate for physical activity and what diet to eat to achieve a certain weight and a more schooled individual may select a more efficient combination of the two. In this context, demand for health inputs derives from the demand for specific health levels (Grossman, Kaestner 1997).

Using the demand for health model of Grossman (1972) reproduced below an individual maximizes an inter-temporal utility function where *i* is the time period, H_0 is the initial health stock, $\gamma_i$is the service flow per unit health stock, and Z_i is total consumption of another commodity:[35]

$$(1)\ U = U(\gamma_0 H_0 \dots \gamma_0 H_n;\ Z_0 \dots Z_n)$$

[35] Grossman's demand for health model is based on Becker's time allocation model (Becker 1965).

The net investment in health stock is defined as gross investment minus depreciation where I_i is gross investment and δ_i is the depreciation rate that may vary with age:

$$(2)\ H_{i+1} - H_i = I_i - \delta_i H_i$$

Individuals generate gross investments in health and the other commodity based on their household production functions where D_i captures health inputs, X_i is the market goods input for the other commodity, TH_i and T_i are the time allocated to produce each commodity, and S_i is the human capital stock:[36]

$$(3)\ I_i = I_i(D_i, TH_i; S_i)$$

$$(4)\ \mathrm{Z_i} = (\mathrm{X_i}, \mathrm{T_i}; \mathrm{S_i})$$

In the gross health investment function (equation 3) market goods may include a variety of health inputs including for instance, medical services, diet, and physical activity. The budget constraint for market goods is:

$$(5) \sum_{\mathrm{i}}^{\mathrm{n}} \frac{\mathrm{P_i D_i + V_i X_i}}{(1+\mathrm{r})^{\mathrm{i}}} = \sum_{\mathrm{i}}^{\mathrm{n}} \frac{\mathrm{W_i TW_i}}{(1+\mathrm{r})^{\mathrm{i}}} + A_0$$

where P_i is the price of health inputs and V_i the price of market goods inputs, W_i is the wage rate, A_0 is initial assets, *r* is the interest rate, and TW_i is hour worked. There is also a time constraint according to which total time available in any given time period, Γ, must equal the time used for all tasks: producing health, producing the other commodity, working, and time lost because of sick days, TL_i:

$$(6)\ TH^i + T^i + TW^i + TL^i = \Gamma$$

Combining the budget constraint and time constraint yields the total wealth constraint:

$$(7) \sum_{i}^{n} \frac{P_i D_i + V_i X_i + \mathrm{W_i(TH_i + T_i + TL_i)}}{(1+r)^i} = \sum_{i}^{\mathrm{n}} \frac{W_i \Gamma}{(1+r)^i} + A_0$$

Individuals then maximize their utility subject to the household production functions and the total wealth constraint. Based on the first-order conditions, the present value of

[36] The production functions are assumed to be homogenous of degree one.

the marginal cost of gross investment must equal the present value of the marginal benefits.

In this model, individual characteristics, in particular schooling, affect the efficiency with which health inputs are combined to generate health outputs. More schooled individuals may be more efficient in producing desired health outputs either because they generate more health outputs from a given set of inputs (productive efficiency), or because they have more and better knowledge of the importance and workings of certain health inputs and therefore choose a better combination of inputs (allocative efficiency) (Grossman 1972).

Another important issue concerning the effect of schooling on health is the possibility of joint production (Grossman, Kaestner 1997). For example, an individual although aware of the consequences, may choose an unhealthy diet and insufficient physical activity if there are other factors that outweigh their negative effect on weight. For instance, by using the time that would have been spent on physical activity on work instead or by choosing a less healthy diet that takes less time to prepare in order to allocate more time to work. However, in general, more schooling is anticipated to improve health outputs.

2.5.2 Estimating the Health Returns to Schooling

The outcome of interest is the BMI of men and women in Indonesia who were 20-50 years old in 2000. The relationship between schooling and BMI is estimated first by ordinary least squares to generate benchmark results then using a two-stage instrumental variable approach to examine the causal effect of schooling.

Schooling and good health are typically positively correlated (Grossman, Kaestner 1997). There are three possible reasons why: more schooling causes better health; better health causes more schooling; or there is no causality but rather unobserved heterogeneity (e.g., innate ability or preferences) causes schooling and health to change in the same direction (Grossman, Kaestner 1997, Schultz 2008).[37]

Due to unobserved heterogeneity, reverse causation, and measurement error, ordinary least squares estimates of the effect of schooling on health tend to be biased and inconsistent (Behrman 1997).[38] To estimate consistently the impact of schooling on

[37] Measurement error is arguably a minor issue as the sample consists of relatively young individuals.
[38] There are several factors that may bias ordinary least squares estimates and the direction of such biases is likely to vary (Behrman 1997).

health one solution is to find a valid instrument, a variable that is highly correlated with the endogenous variable (relevance) and that is not correlated with the error term in the outcome equation (exogeneity). This chapter uses exposure to the SD INPRES programme and father's schooling, as discussed earlier, to capture exogenous variation in schooling. The models estimate the total effect of schooling but do not reveal the channels, for instance, changes in diet and physical activity, through which schooling influences BMI outcomes.

Benchmark Model

The demand for health (BMI) equation is used to estimate the relationship between schooling and BMI. It controls for individuals' marital status, rural/urban location, and household assets. The equation also includes dummies to control for district and year of birth specific effects. Interactions of year of birth with primary enrollment prior to the SD INPRES programme and with the water supply and sanitation programme in each district in 1971 are included to account for the targeted allocation of the SD INPRES programme and for other public programmes that may have affected health respectively (Duflo 2001). The benchmark BMI equation is then:

$$BMI_{ijk} = c_0 + \alpha_j + \beta_k + \gamma_1 S_{ijk} + \gamma_2 A_i + \gamma_3 M_i + \gamma_4 U_i + \gamma_5 C_j + u_{ijk} \tag{8}$$

i=individual
j=district of birth
k=year of birth
c_0 = constant
α_j = district of birth specific effect
β_k = year of birth specific effect
S_{ijk} = completed years of schooling
A_i= household assets
M_i= marital status
U_i= urban location
C_j = interactions between year of birth and primary enrollment rate and between year of birth and water supply and sanitation programme in 1971 in each district
u_{ijk}=unobservables (e.g., preferences)

Instrumental Variable Estimation

To assess the causal effect of schooling on BMI a two-stage instrumental variable approach is used. In the first-stage, schooling is regressed on all exogenous variables including SD INPRES programme exposure and father's schooling to generate the fitted values for the second-stage BMI regression. The first-stage schooling equation is given by:

$$(9) \quad S_{ijk} = c_0 + \alpha_j + \beta_k + \delta_1 E_i + \delta_2 P_{ijk} + \delta_3 A_i + \delta_4 M_i + \delta_5 U_i + \delta_6 C_j + \varepsilon_{ijk}$$

E_i = completed years of schooling of father
P_{ijk} = programme exposure: cumulative number of schools constructed per 1,000 children aged 5-14 in district of schooling in year prior to starting primary school
ε_{ijk}=unobservables (e.g., innate ability)
All other variables defined as above

The second-stage equation for the impact of schooling on BMI is estimated using the fitted values of schooling from the first-stage. The identification comes from E_i and P_{ijk} in the first-stage. In the BMI equation, γ_1 captures the causal effect of schooling on BMI:

$$(10) \quad BMI_{ijk} = c_0 + \alpha_j + \beta_k + \gamma_1 \hat{S}_{ijk} + \gamma_2 A_i + \gamma_3 M_i + \gamma_4 U_i + \gamma_5 C_j + \upsilon_{ijk}$$

$\hat{S}_{ij}$= predicted schooling
υ_{ijk}=unobservables (e.g., preferences)
All other variables defined as above

Exposure to the SD INPRES programme varies at the district level. To allow for correlation in the error structure for individuals within districts all standard errors are clustered at the district level.

2.6 Results and Discussion

This section first presents the ordinary least squares benchmark estimates. It then discusses the first-stage schooling regressions and tests of instrument relevance and validity. Then the findings on the causal effect of schooling on BMI for men and women in Indonesia are discussed. BMI equations are also estimated for the sub-sample of men for whom data on occupation type are available in an attempt to assess if the causal effect of schooling on overweight in Indonesia is partly working through occupation type. For ease of exposition, throughout the section the results for women are presented in parentheses next to those for men.

2.6.1 Results from the Benchmark Model

The results for the BMI equation estimated by ordinary least squares are shown in columns 1 and 3 in Table 23 for the samples of 4,300 men and 4,821 women. These are estimates of the total effect of schooling on BMI, which do not reveal the channels through which schooling may influence BMI.

Treating schooling as exogenous it is significantly related with higher BMI for men and a one standard deviation rise in schooling is associated with a 0.53 higher BMI. However, there is no significant effect of schooling for women. This compares to existing findings for Indonesia where men who have completed junior secondary school or more have significantly higher BMI than those with less schooling but is different for women whose BMI is found to rise with increasing schooling up until completion of primary schooling but with no additional effect thereafter (Witloear, Strauss et al. 2011).

Wealth affects BMI in the expected direction for men (women): those in the poorest tertile and middle tertile respectively have significantly lower BMI than those in the richest tertile similar to existing studies (Monteiro, Moura et al. 2004, Witloear, Strauss et al. 2011). For a man (woman) moving from the richest to the middle tertile is associated with a 0.54 (0.49) lower BMI and from the richest to the poorest tertile with a 0.65 (0.77) lower BMI.

Living in an urban area and being married are both positively and significantly related to BMI, and the effects are much larger for women. On average, for men (women) in urban areas BMI is 0.38 (0.95) higher and for married men 0.54 (1.18) higher. This is in line with existing evidence that overweight is more common in urban areas and that married people are more likely to be overweight (Averett, Sikora et al. 2008, Popkin 1999).

2.6.2 Instrumental Variable Results

The results from the first-stage schooling regressions are presented in Table 22. SD INPRES programme exposure and father’s schooling, the excluded instruments, are highly significant and positively related to schooling both for men and for women but the effect of programme exposure is smaller for women. For men (women) a one standard deviation increase in SD INPRES programme exposure is associated with 0.51 (0.26) years of additional schooling and a one standard deviation increase in father’s schooling with 1.3 (1.4) additional years of schooling. The F-statistic for joint significance of the excluded instruments is high, 151 for the male sample and 274 for the female sample respectively, indicating that the instruments are highly relevant.

Poorer men and women have significantly less schooling as anticipated and the effects are of similar magnitude. Men (women) in the poorest tertile have on average 2.3 (2.3) years less schooling and men (women) in the middle tertile 1.3 (1.3) years less schooling than men (women) in the richest tertile.

Being married is associated with less schooling for both men and women but only significantly so for men who have 0.3 years less schooling than their unmarried counterparts. Men and women living in an urban area have significantly more schooling. Men (women) in urban areas have about 0.6 (0.7) additional years of schooling compared to those living in rural areas.

The null of schooling exogeneity is rejected for all models, suggesting that schooling is endogenous in this context.[39] Hansen's J test for overidentification of all instruments does not reject its null, which indicates that the instrumental variables are valid providing further confidence for the choice of instruments.[40] An equation for BMI that adds father's schooling as an independent variable is also estimated and the coefficient on father's schooling is not significantly different from zero supporting the argument that father's schooling does not directly influence BMI (results not reported).

The results from the second-stage BMI estimations are shown in Table 23. More schooling causes higher BMI for men in Indonesia (column 2). A one standard deviation (4 years) increase in schooling causes a 0.73 increase in BMI. For example, for a man of mean height (1.62 metres) and mean BMI (21.2), four additional years of schooling would increase his weight by 2 kilograms. By contrast, schooling does not cause BMI for women (column 4).

These results can be compared to the study for Germany, which finds that more schooling leads to a higher probability of being obese for men but finds no significant causal effect of schooling on BMI for women (Jürges, Reinhold et al. 2011). Moreover, one of the studies for the United States finds that for women obesity differences by schooling are entirely the result of selection into different schooling levels (Conti, Heckman 2010).

The instrumental variable effects of schooling on BMI are significantly larger than the ordinary least squares estimates: 0.183 compared to 0.132 respectively for men. That the instrumental variable estimates are notably larger is consistent with the findings of other studies (Arendt 2005, Grabner 2009, Clark, Royer 2010), and may be

[39] P-value = 0.0000 for all cases.
[40] P-value = 0.46 for men, P-value = 0.96 for women.

due to measurement error (Griliches 1977), or heterogeneous returns to schooling (Imbens, Angrist 1994). In the Indonesian case, those who decided to attend primary school in response to the increase in the primary school supply may have relatively higher health returns to schooling than those whose schooling decisions were not affected by the programme in which case the instrumental variable estimator gives the local average treatment effect. Nevertheless, since the SD INPRES programme affected such a large proportion of the sample the local average treatment effect should approach the average treatment effect (Devereux, Fan 2011).

All other results are highly similar across estimation methods. Men and women who are poorer have significantly lower BMI and men and women who are married or live in an urban setting have significantly higher BMI.

To address the issue of over-age children discussed in Section 2.4 all models are also estimated using female and male samples that include children who were 8-12 years old (over-age) in the year SD INPRES schools were constructed in their district and which treats them as fully exposed to the programme and all results are highly similar (results not reported).

To explore where along the BMI distribution schooling has an impact, i.e., whether more schooling moves men into the overweight (underweight) category, probit models treating schooling as exogenous and endogenous respectively, are estimated with an indicator variable equal to one if BMI categorises an individual as overweight (underweight) and zero otherwise.[41] For the sample consisting of 3,287 normal and overweight men, those with more schooling are significantly more likely to be overweight (Table 24). Computing the average marginal effect, one additional year of schooling leads to an average increase of 0.115 in the probability of being overweight. Further, living in an urban as opposed to a rural area increases the probability of being overweight by 0.246. For the 3,618 men in the sample composed of normal and underweight men, those with more schooling are significantly less likely to be underweight in the regressions treating schooling as exogenous but there is no significant relationship when instrumenting for schooling (Table 24).

For a preliminary exploration of potential channels through which schooling may cause higher BMI among men a model that includes dummy variables for occupation type: blue collar (agriculture, mining, construction, manufacturing, and utilities), service

[41] The instrumental variables are valid according to the F test of excluded instruments and Hansen J statistic respectively (first stage regression results not shown).

industry (services, communications, and transportation), and white collar (finance, business, and social services) are estimated. Summary statistics for this sub-sample of 3,883 men are presented in Table 19. The men who do not report their occupation have less exposure to the SD INPRES programme, their fathers have less schooling, and they are less likely to be married, which may affect the comparability of results. Therefore, the main model is first re-estimated for the sub-sample and then with the additional occupation variables.

Table 25 shows the first-stage schooling regressions and Table 26 the ordinary least squares and second-stage BMI regressions for the sub-sample of men for whom occupation data are available. The first-stage schooling and second-stage BMI results are generally similar to those for the full sample. Comparing the model without occupation variables (column 2) to the one that controls for occupation type (column 4) the causal effect of schooling on BMI declines somewhat from 0.78 to 0.72.

As above, models using an indicator variable equal to one if a man is overweight (underweight) and zero otherwise as the dependent variables are estimated but now including the occupation variables.[42] As before men with more (less) schooling are significantly more likely to be overweight (underweight) (Table 27). Men working in services are significantly more likely to be overweight than men with blue collar jobs and so are men with white collar jobs in the regressions treating schooling as exogenous. Based on the calculated average marginal effect working in services compared to having a blue collar job raises the probability of being overweight by 0.284.

This provides some very preliminary support that one mechanism through which more schooling causes higher BMI for men is the general shift from relatively energy-intensive manual work to more sedentary service industry and perhaps also white collar jobs. At a minimum, this result precludes that more schooling reduces BMI for men. There is some preliminary evidence of such a mechanism in other countries as well (Jürges, Reinhold et al. 2011).

More schooling may not only cause higher BMI via a shift from relatively energy-intensive manual jobs to more sedentary ones but also through longer hours works, which leads to less time for physical activities outside of work and may induce a change from meals prepared at home to less healthy pre-prepared and fast food (Courtemanche

[42] The instrumental variables are again valid according to the F test of excluded instruments and Hansen J statistic respectively (first stage regression results not shown).

2009). New evidence for Indonesia in addition finds that more schooling causes higher wages and that men exposed to the SD INPRES programme have a higher relative risk of being in a white collar than a blue collar occupation and a higher probability of having a wage earning job than being self-employed (Chapter 3). This means that higher BMI may also be linked to higher wages that allow individuals a higher caloric intake with potential implications for BMI if there is no adjustment to calories expended and also by the shift away from blue collar jobs.

2.7 Conclusion

This chapter is the first to assess the causal effect of schooling on BMI in a developing country. Two results stand out for Indonesia: more schooling causes higher BMI for men and there is no causal effect of schooling on BMI for women. Such a differential effect of schooling on BMI of men and women has previously been found in developed countries (Conti, Heckman 2010). These results compare to the findings of a recent non-causal study of Indonesia, which finds that men with more schooling have significantly higher BMI than men with no schooling, and that women with more schooling have significantly higher BMI but that there is no additional effect after completion of junior secondary school (Witloear, Strauss et al. 2011).

This chapter also provides some very preliminary evidence that the shift from blue collar, relatively energy-intensive jobs to relatively more sedentary white collar and service sector jobs may be one contributing factor to why more schooling increases BMI for men. Higher wages related to more schooling may also contribute by allowing a higher calorie diet.

However, new research is required to understand through which other mechanisms, for instance, changes in social norms and the role of deprived backgrounds, schooling causes higher BMI (Graham, Felton 2005, Akee, Simeonova et al. 2010). Such information is crucial for the design of public policies that can effectively decelerate the increase in overweight and overweight-related chronic diseases in Indonesia and other developing countries.

Table 15. Literacy, primary enrollment, and schooling attainment in Indonesia, 1971

	Total	Female	Male	Rural	Urban
Literacy (%)	60	49	71	55	79
Population with no	41	52	30	45	22
Population with complete	19	15	24	27	18

Source: (World Bank 1975)

Table 16. SD INPRES programme investments, 1973/74-1978/79

	SD INPRES programme investments			
	New primary schools	Primary school books (millions)	SD INPRES allocation (billions of rupiah)	SD INPRES allocation as % of total public spending
1973/74	6,000	6.6	17.2	1.5
1974/75	6,000	6.9	19.7	1.0
1975/76	10,000	7.3	49.9	1.9
1976/77	10,000	8.6	57.3	1.6
1977/78	15,000	7.3	85.0	2.0
1978/79	15,000	8.5	111.8	2.1
Total	62,000	45.2	341	-

Note: All schools constructed had 3-classroom blocks. Textbooks in mathematics and Bahasa Indonesia.

Sources: (World Bank 1989, World Bank 1984).

Table 17. Variable description

Variable	Description
asset	Weighted household asset index
blue collar	Dummy variable equal to 1 if respondent has blue collar job
bmi	Body mass index for each respondent (weight in kg/height in metres squared)
district	Dummy variable for respondent's district of schooling
father's schooling	Father's years of completed schooling
married	Dummy variable equal to 1 if respondent is currently married
middle tertile	Dummy variable equal to 1 if respondent is in middle third based on asset index
poorest tertile	Dummy variable equal to 1 if respondent is in poorest third based on asset index
proportion enrolled in 1971	Fraction of children enrolled in primary school in each district in 1971
richest tertile	Dummy variable equal to 1 if respondent is in richest third based on asset index
schooling	Years of completed schooling
SD INPRES pgm exposure	Exposure to SD INPRES school construction programme
service industry	Dummy variable equal to 1 if respondent has service industry job
urban	Dummy variable equal to 1 if urban
water supply and sanitation	Water supply and sanitation programme per capita spending in each district 1973-1978
white collar	Dummy variable equal to 1 if respondent has white collar job
year of birth	Dummy variable for respondent's year of birth

Table 18. Summary statistics (full sample)

	Men		*Women*	
	Mean	Std. Dev.	Mean	Std. Dev.
asset	6.0	5.7	6.1	5.8
bmi	21.2	3.1	22.2	3.7
father's schooling	5.0	3.8	4.9	3.9
married	0.68	0.47	0.77	0.42
proportion enrolled in 1971	0.17	0.06	0.16	0.06
schooling	8.5	4.0	7.5	4.3
SD INPRES pgm exposure	0.92	0.96	0.93	0.96
urban	0.49	0.50	0.48	0.50
water supply and sanitation	0.45	0.22	0.45	0.20
year of birth	1968	9	1968	9
observations	4300		4821	

Table 19. Summary statistics (small sample)

	Men	
	Mean	Std. Dev.
asset	5.9	5.6
blue collar	0.55	0.50
bmi	21.3	3.1
father's schooling	4.8	3.7
married	0.74	0.4
proportion enrolled in 1971	0.16	0.06
schooling	8.3	4.0
SD INPRES pgm exposure	0.86	0.94
service industry	0.23	0.42
urban	0.47	0.50
water supply and sanitation	0.45	0.2
white collar	0.22	0.41
year of birth	1967	9.3
observations	3883	

Table 20. Mean schooling, father's schooling, and BMI by gender and SD INPRES programme exposure

	SD INPRES programme		
		Men	
Mean	not exposed	exposed	difference
schooling	6.8	9.3	2.4*
father's schooling	3.7	5.7	2.0*
BMI	22.1	20.8	-1.2*
observations	1404	2896	
		Women	
Mean	not exposed	exposed	difference
schooling	5.1	8.7	3.6*
father's schooling	3.6	5.6	2.0*
BMI	23.3	21.7	-1.6*
observations	1519	3302	

Note: *statistically different at one percent.

Table 21. Frequency of BMI outcomes by gender

	Men (%)	Women (%)
underweight	15.9	13.6
normal weight	72.8	65.4
overweight or obese	11.3	21.0
Total	100	100

Table 22. First-stage schooling regressions (full sample)

	Men	*Women*
	(1)	(2)
Dependent variable: schooling	OLS	OLS
poorest tertile	-2.2887***	-2.3293***
	(0.1564)	(0.1418)
middle tertile	-1.2757***	-1.3074***
	(0.1517)	(0.1312)
married	-0.3318**	-0.1096
	(0.1317)	(0.1160)
urban	0.5914***	0.6867***
	(0.1869)	(0.1489)
SD INPRES pgm exposure	0.5266***	0.2691***
	(0.1504)	(0.0912)
father's schooling	0.3473***	0.3714***
	(0.0201)	(0.0161)
district	Y	Y
year of birth	Y	Y
year of birth*proportion enrolled 1971	Y	Y
year of birth*water supply and sanitation	Y	Y
F test of excluded instruments (F statistic/P-value)	151.42/0.000	274.29/0.000
Observations	4300	4821
R-squared	0.240	0.287

Robust standard errors clustered at district level in parentheses. ***p<0.01, **p<0.05, *p<0.1. All regressions include a constant. Excluded category is richest tertile.

Table 23. BMI regressions (full sample)

	Men		*Women*	
	(1)	(2)	(3)	(4)
Dependent variable: body mass index	OLS	IV	OLS	IV
schooling	0.1318***	0.1833***	0.0006	0.0079
	(0.0155)	(0.0483)	(0.0174)	(0.0423)
poorest tertile	-0.6502***	-0.4970***	-0.7727***	-0.7493***
	(0.1250)	(0.1845)	(0.1678)	(0.1889)
middle tertile	-0.5425***	-0.4579***	-0.4933***	-0.4804***
	(0.1337)	(0.1560)	(0.1627)	(0.1667)
married	0.5351***	0.5501***	1.1773***	1.1788***
	(0.1222)	(0.1170)	(0.1273)	(0.1241)
urban	0.3844***	0.3352**	0.9456***	0.9377***
	(0.1441)	(0.1430)	(0.1807)	(0.1848)
district	Y	Y	Y	Y
year of birth	Y	Y	Y	Y
year of birth*proportion enrolled 1971	Y	Y	Y	Y
year of birth*water supply and sanitation	Y	Y	Y	Y
Hansen J statistic (overidentification test of all instruments, P-value)		0.459		0.956
Observations	4300	4300	4821	4821
R-squared	0.180	0.060	0.173	0.052

Robust standard errors clustered at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

All regressions include a constant term. Schooling is instrumented. Excluded category is richest tertile.

Table 24. Over- and underweight regressions (men)

	Overweight		Underweight	
Dependent variable:	(1)	(2)	(3)	(4)
Overweight/underweight dummy	Probit	IV Probit	Probit	IV Probit
schooling	0.0602***	0.1148***	-0.0235**	0.0177
	(0.009)	(0.027)	(0.010)	(0.028)
poorest tertile	-0.3975***	-0.2221*	-0.0078	0.1119
	(0.087)	(0.119)	(0.093)	(0.133)
middle tertile	-0.3586***	-0.2595**	-0.0511	0.0164
	(0.105)	(0.113)	(0.097)	(0.116)
married	0.1261	0.1465	-0.2270***	-0.2120***
	(0.113)	(0.109)	(0.075)	(0.075)
urban	0.2968***	0.2456***	0.0394	0.0014
	(0.090)	(0.093)	(0.080)	(0.083)
district	Y	Y	Y	Y
year of birth	Y	Y	Y	Y
year of birth*proportion enrolled 1971	Y	Y	Y	Y
year of birth*water supply and sanitation	Y	Y	Y	Y
Hansen J statistic (overidentification test of all instruments, P-value)		0.320		0.421
Observations	3287	3287	3618	3618

Robust standard errors clustered at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term. Schooling is instrumented. Excluded asset category is richest tertile.

Table 25. First-stage schooling regressions (small sample, men)

	Men	
	(1)	(2)
Dependent variable: schooling	OLS	OLS
poorest tertile	-2.3066***	-2.0843***
	(0.1620)	(0.1616)
middle tertile	-1.2803***	-1.1456***
	(0.1590)	(0.1597)
married	-0.3477**	-0.2793**
	(0.1358)	(0.1314)
urban	0.5259***	0.3786*
	(0.1960)	(0.1955)
service industry		0.2265
		(0.1467)
white collar		1.6688***
		(0.1920)
SD INPRES pgm exposure	0.5575***	0.4932***
	(0.1795)	(0.1606)
father's schooling	0.3499***	0.3310***
	(0.0212)	(0.0217)
district	Y	Y
year of birth	Y	Y
year of birth*proportion enrolled 1971	Y	Y
year of birth*water supply and sanitation	Y	Y
F test of excluded instruments (F statistic/P-value)	136.44/0.000	117.32/0.000
Observations	3883	3883
R-squared	0.239	0.271

Robust standard errors clustered at district level in parentheses. ***p<0.01, **p<0.05, *p<0.1. All regressions include a constant. Excluded asset category is richest tertile, excluded occupation category is blue collar.

Table 26. BMI regressions (small sample, men)

	Men			
	(1)	(2)	(3)	(4)
Dependent variable: body mass index	OLS	IV	OLS	IV
schooling	0.1366***	0.1944***	0.1286***	0.1824***
	(0.0159)	(0.0483)	(0.0160)	(0.0508)
service industry			0.4962***	0.4769***
			(0.1256)	(0.1211)
white collar			0.3845**	0.2788*
			(0.1488)	(0.1672)
poorest tertile	-0.7149***	-0.5433***	-0.6513***	-0.5082***
	(0.1334)	(0.1905)	(0.1344)	(0.1866)
middle tertile	-0.5592***	-0.4639***	-0.5247***	-0.4459***
	(0.1354)	(0.1571)	(0.1364)	(0.1555)
married	0.5548***	0.5703***	0.5422***	0.5529***
	(0.1253)	(0.1193)	(0.1238)	(0.1181)
urban	0.3678**	0.3182**	0.3060**	0.2707*
	(0.1518)	(0.1493)	(0.1473)	(0.1446)
district	Y	Y	Y	Y
year of birth	Y	Y	Y	Y
year of birth*proportion enrolled 1971	Y	Y	Y	Y
year of birth*water supply and sanitation	Y	Y	Y	Y
Hansen J stat (overidentification test of all		0.312		0.355
Observations	3883	3883	3883	3883
R-squared	0.191	0.066	0.196	0.072

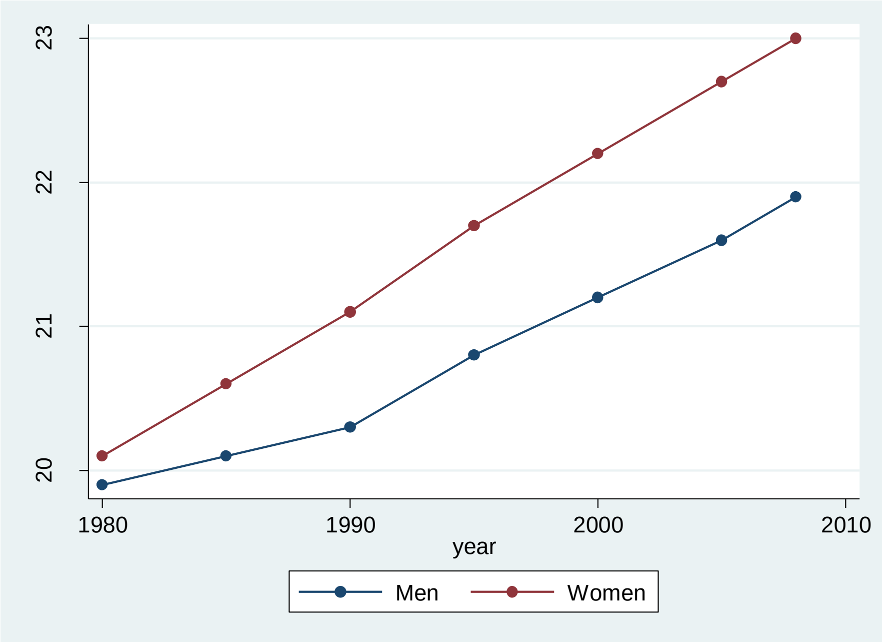
Robust standard errors clustered at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term. Schooling is instrumented. Excluded asset category is richest tertile, excluded occupation category is blue collar.

Table 27. Over- and underweight regressions (small sample, men)

	Overweight		*Underweight*	
Dependent variable:	(1)	(2)	(3)	(4)
Overweight/underweight dummy	Probit	IV Probit	Probit	IV Probit
schooling	0.0627***	0.1141***	-0.0196**	0.0200
	(0.011)	(0.030)	(0.010)	(0.031)
poorest tertile	-0.3605***	-0.2125*	0.0361	0.1377
	(0.086)	(0.119)	(0.102)	(0.140)
middle tertile	-0.3398***	-0.2551**	-0.0322	0.0253
	(0.108)	(0.120)	(0.108)	(0.124)
service industry	0.3101***	0.2835***	-0.1443**	-0.1632**
	(0.081)	(0.081)	(0.070)	(0.071)
white collar	0.1982**	0.0764	-0.0251	-0.0958
	(0.092)	(0.112)	(0.084)	(0.092)
married	0.1823*	0.1917*	-0.2102**	-0.2015**
	(0.111)	(0.107)	(0.082)	(0.082)
urban	0.2869***	0.2545***	0.0473	0.0231
	(0.091)	(0.094)	(0.085)	(0.087)
district	Y	Y	Y	Y
year of birth	Y	Y	Y	Y
year of birth*proportion enrolled 1971	Y	Y	Y	Y
year of birth*water supply and sanitation	Y	Y	Y	Y
Hansen J stat (overidentification test of all instruments, p-value)		0.482		0.506
Observations	2967	2967	3213	3213

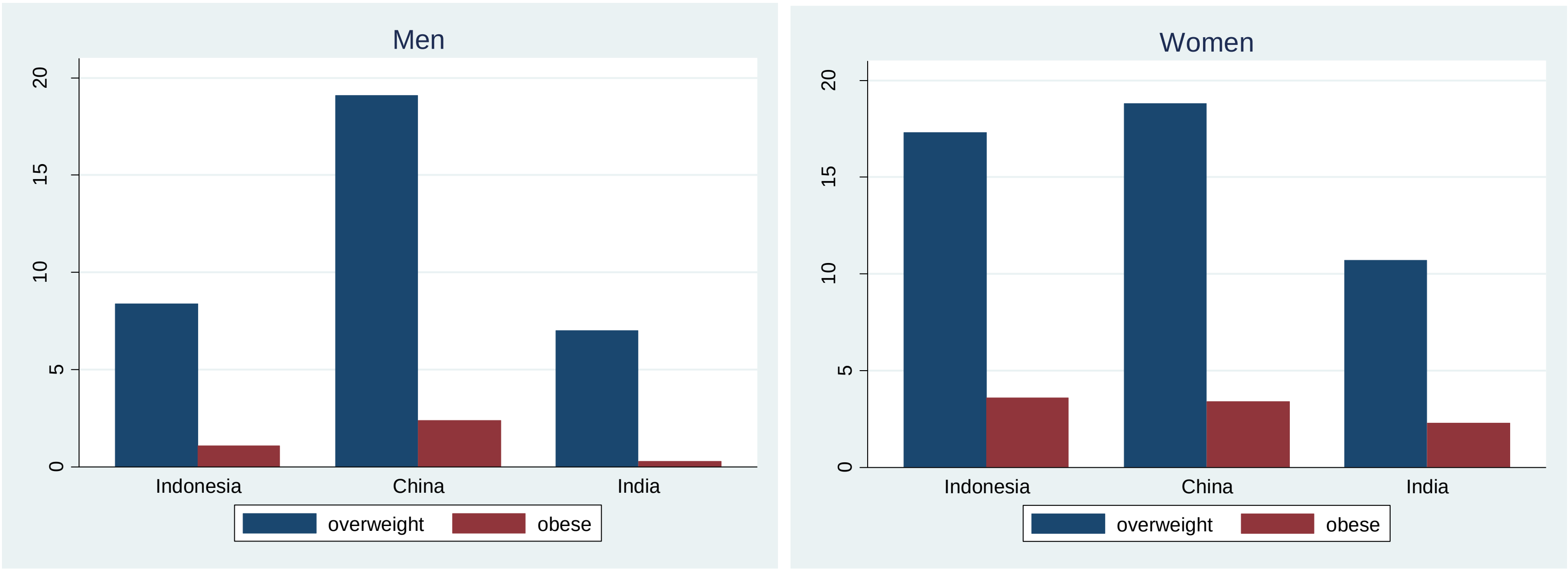
Robust standard errors clustered at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term. Schooling is instrumented. Excluded asset category is richest tertile, excluded occupation category is blue collar.

Figure 8. Mean BMI in Indonesia by gender, 1980-2008



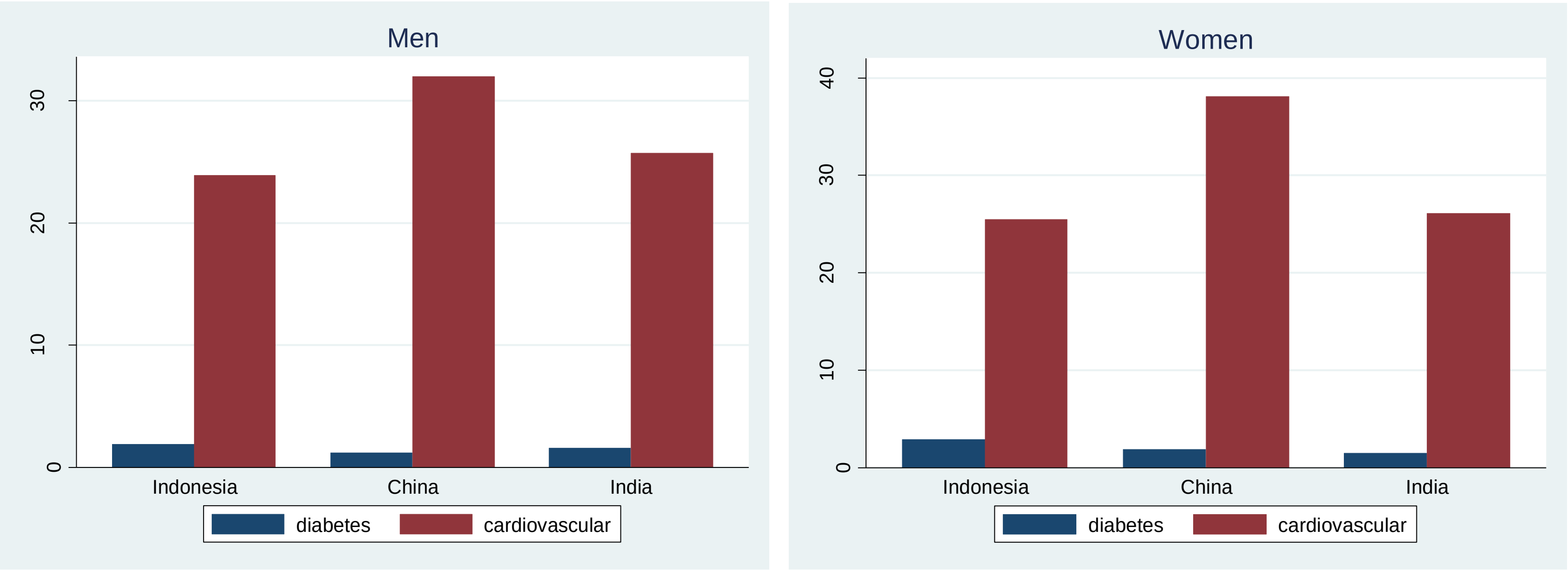
Source: (WHO 2011b).

Figure 9. Overweight in Indonesia, China, and India by gender, 2004



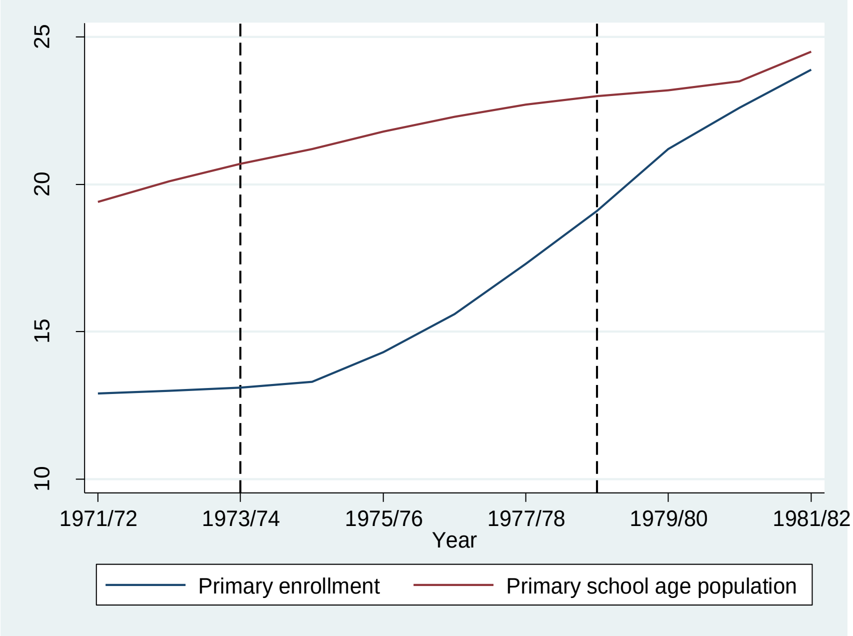
Source: (WHO 2011a).

Figure 10. Overweight related causes of death in Indonesia, China, and India by gender, 2004



Source: (WHO 2011b).

Figure 11. Primary enrollment and school-age population, 1971-1982



Note: School-age population is children aged 7-12 years.
Source: (World Bank 1975, World Bank 1989).

Figure 12. Empirical CDF of BMI by gender, 2000

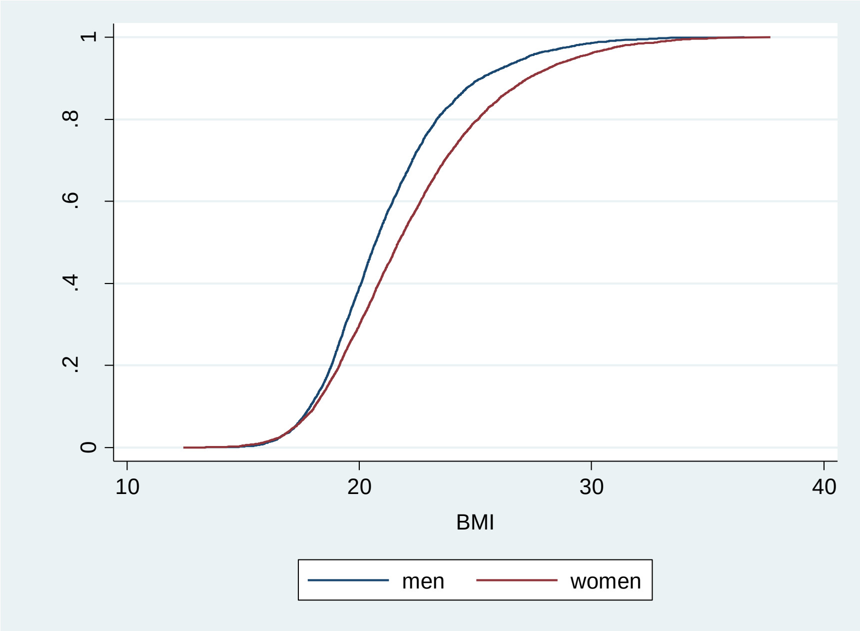
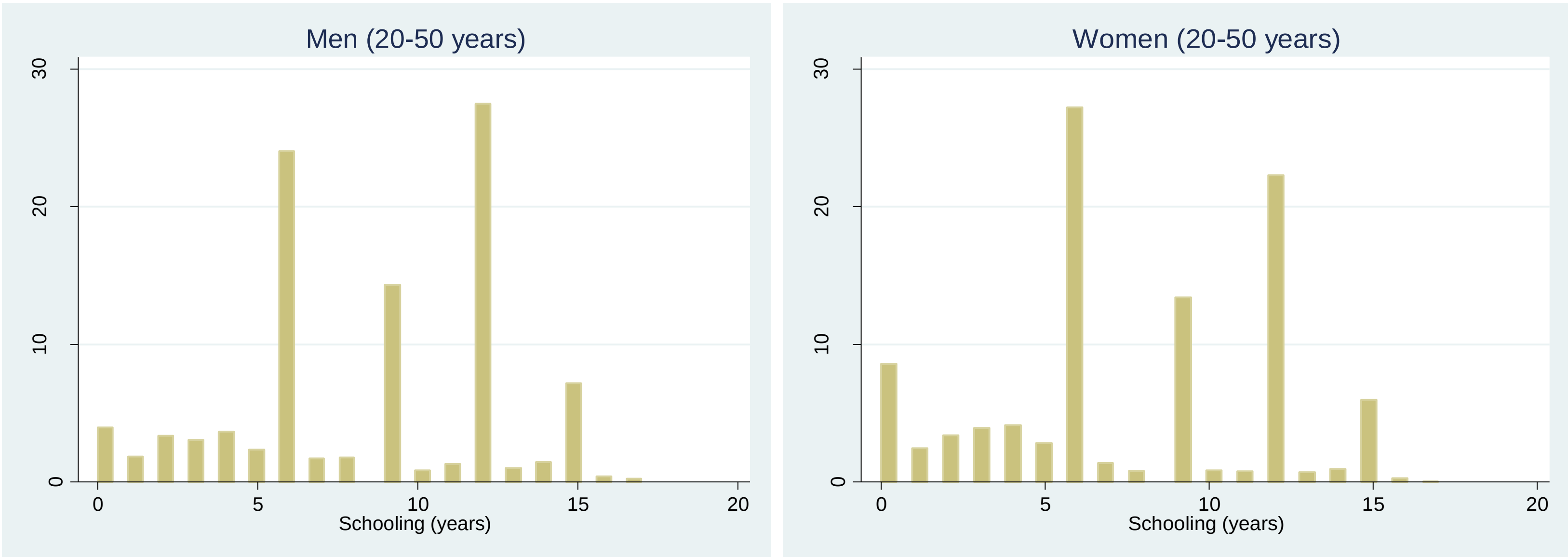


Figure 13. Distribution of schooling by gender, 2000



Chapter 3: Do Supply-Side Education Programmes Work? The Impact of Increased School Supply on Schooling and Wages in Indonesia Revisited

Abstract

Indonesia each year allocates a large proportion of its total public spending to education and it is important to understand whether different groups, for instance, children from less advantageous socioeconomic backgrounds or girls benefit differentially from these public investments. It is also desirable to comprehend whether schooling translates into increases in wages that are similar in size for both for men and for women who obtain additional schooling. This chapter uses the large-scale Presidential Instruction Primary School construction programme (SD INPRES) rolled out in Indonesia in the 1970s to examine the effect of increased school supply on schooling attainment: overall, by gender, and by socioeconomic background. It also constructs a new SD INPRES programme exposure variable that is used as an instrument for schooling to assess the causal effect of schooling on wages and whether the additional schooling induced by the programme had differential impacts for men and women. To preview the findings, SD INPRES programme exposure significantly increased schooling both for men and for women. Moreover, women benefitted more from the SD INPRES programme than men and so did individuals from less advantageous socioeconomic backgrounds contributing to a narrowing of schooling gaps by gender and by socioeconomic background. In addition, more schooling is found to cause higher wages for men and women and there appears to be an additional positive effect on wages for women through the additional schooling induced by the SD INPRES programme. Further, men exposed to the SD INPRES programme have a significantly higher relative risk of having a white collar than blue collar occupation and a higher probability of being in wage employment rather than self-employed which suggests that more schooling partly causes higher wages through occupation choice and employment type.

3.1 Introduction

Individuals with more schooling tend to earn more, have better health outcomes, lower fertility, and live longer (Card 1999, Grossman, Kaestner 1997, Strauss, Thomas 1995). However, schooling can also provide benefits to society by promoting economic growth and development, narrowing gender earnings gaps, increasing earnings mobility, and encouraging civic participation (Barro 2001, Behrman, Stacey 1997, Fields 2008, Milligan, Moretti et al. 2004, Orazem, King 2008).

The private and social benefits of schooling are the reason governments each year allocate large proportions of their total public spending to education. However, despite large-scale investments in schooling, gender gaps in schooling persist and so do differences in schooling by socioeconomic background (Orazem, King 2008).

Indonesia also allocates a large proportion of its total public spending to education and it is crucial to understand if different groups, for instance, children from less advantageous socioeconomic backgrounds or girls benefit differentially from these public investments.[43] It is also important to comprehend whether schooling translates into increases in wages that are similar in size for both men and women who obtain the additional schooling.[44] Existing empirical studies of the causal effect of schooling on the return to schooling tend to focus, however, on the overall effect on wages (Card 1999, Card 2001).

This chapter empirically examines the effect of the large-scale Presidential Instruction Primary School construction programme (SD INPRES) in Indonesia in the early 1970s on schooling attainment and wages with the objective of answering three questions. First, did the SD INPRES programme raise schooling attainment? Second, did the SD INPRES programme have a differential effect on schooling attainment depending on socioeconomic background or gender? Third, did more schooling cause higher wages and if so was the effect the same across men and women?

The SD INPRES programme has previously been used to assess the effect of increased school supply on schooling attainment and the return to schooling in a seminal paper (Duflo 2001). However, the effect of the SD INPRES programme was

[43] In Indonesia, education accounted for 18 percent of total public spending in 2006 and for 12 percent in 1975 (Herold, Khan et al. 1983, UNESCO 2008).
[44] Assuming they are similar with respect to all other relevant characteristics.

only examined for men and potential differential effects by socioeconomic background were not explored. The programme has also been used to examine the effect of increased school supply on schooling attainment (Hertz, Jayasundera 2007). Nevertheless, that study does not examine the effect of schooling on wages and uses a different SD INPRES programme exposure variable, which produces inconclusive results.

This chapter thus aims to pull all the pieces together in order to provide results on the effect of the SD INPRES programme on schooling: overall, by gender, and by socioeconomic background, and on the causal effect of schooling on wages with a particular focus on whether the additional schooling induced by the programme had differential impacts for men and women.

To preview the findings, SD INPRES programme exposure significantly increased schooling both for men and for women. Moreover, women benefitted more from the SD INPRES programme than men and so did individuals from less advantageous socioeconomic backgrounds thus contributing to a narrowing of schooling gaps by gender and by socioeconomic background. More schooling is found to cause higher wages for men and women but there appears to be an additional positive effect on wages for women through the additional schooling induced by the SD INPRES programme, which contributes towards reducing the gender wage gap.

The remainder of this chapter is organized as follows. The next section reviews the literature relevant to the analysis in this chapter. Section 3.3 describes the primary school system in Indonesia prior to the SD INPRES school construction programme and the programme itself. In Section 3.4 the sample and data used to assess the effect of increased school supply on schooling and whether more schooling causes higher wages in Indonesia are discussed. The conceptual framework and empirical methodology on which the estimations are based are outlined in Section 3.5. In Section 3.6 the findings and robustness checks are presented and discussed. Finally, Section 3.7 concludes.

3.2 Literature Review

The literature on the relationship between schooling and wages is vast (Card 1999, 2001, Strauss, Thomas 1995, Schultz 1988). To examine the causal effect of schooling on wages instrumental variable methods are frequently used because schooling is arguably endogenous due to omitted variables and measurement error, which leads to bias in standard ordinary least squares estimates (Ashenfelter, Harmon et al. 1999, Card 1999). A valid instrumental variable must be highly correlated with the endogenous schooling variable (relevance) and must not affect the outcome of interest, wages, directly (exogeneity). Much of the recent literature has used changes in institutional features of schooling systems such as minimum school leaving age and distance to school as instruments for schooling, which are argued to be plausibly exogenous (Card 2001). The existing evidence on the causal effect of schooling on wages, most of which is for developed countries, finds that more schooling causes higher wages (Ashenfelter, Harmon et al. 1999, Card 1999, Schultz 1988).

The focus in this section is on studies for developing countries that use measures of schooling availability as instrumental variables to assess the causal effect of schooling on wages. The first-stage regressions in these studies are of interest in their own right for understanding the effect of supply-side school infrastructure policies on school attainment. One study on Indonesia that examines the relationship between the expansion of school supply and schooling attainment but not the effect of schooling on wages is also reviewed.

A study of the return to primary schooling for men in Honduras uses variation in schooling availability measured by the number of primary school teachers per capita as an instrument for schooling (Bedi, Gaston 1999). The argument is that children who were school-age at times when schooling availability was higher faced lower costs of schooling and therfore would obtain more schooling than those who were of school-age when schooling availability was relatively low. The sample consists of 2,014 men aged 16-64 years. The ordinary least squares estimates of the return to schooling treating schooling as exogenous suggest a return to schooling of 6.1 percent, using an alternative sample and adjusted specification the estimated return is 14.0 percent. Parental schooling is found to be positively and significantly related to wages, raising earnings by about 2.5 percent and this effect it is argued to work through ability. In the first-stage schooling regressions, higher schooling availability is significantly associated with more schooling and the F-test for excluded instruments suggests that the instrument is relevant. Moreover, the children of parents with more schooling have significantly more schooling themselves. The instrumental variable estimates of the effect of schooling are roughly 2.5 times larger than the ordinary least squares estimates and the estimated return to schooling is 16.9 percent. However, parental schooling is not longer significantly related to the return to schooling. There are potential issues with the chosen measure of schooling availability, for example, if the number of primary teachers per capita is higher in richer schools, which may cater to students that are systematically different from students in poorer schools along unobserved characteristics.

Using distance to secondary school and its square and the presence of private secondary schools as instruments for schooling the return to schooling is estimated for rural Philippines (Maluccio 1998). Several household characteristics at the time the individuals in the sample attended school are used as additional instruments including father's and mother's schooling and household wealth variables. However, the results from the models that include these additional instruments are not discussed here as these instruments may not be valid (Dearden 1999). The sample consists of men and women who were 20-44 years old in 1994. The dependent variable is daily wages and according to the ordinary least squares estimates the return to schooling is 7.3 percent. In the first-stage schooling regressions, distance to the closest secondary school is negatively and significantly related to schooling and the F-statistic for the excluded instruments ranges from 8.2 to 15.6 indicating that the instruments are relevant. Given the potential problems related to using parental schooling as instrumental variables, only the results based on distance and squared distance to secondary school and presence of private secondary schools are discussed. The returns to schooling is 14.5 percent based on this set of instrumental variables compared to the ordinary least squares estimate of 7.3 percent. It is argued that about a third the difference in the estimates is due to measurement error and the remainder to heterogenous returns to schooling.

The Presidential Instruction Primary School construction programme (SD INPRES) used in this chapter has previously been exploited to assess the effect of increased school supply on schooling attainment and the return to schooling in a seminal paper on Indonesia (Duflo 2001). Here, the main features of the earlier study and the results most relevant to the findings in this chapter are reviewed whereas the SD INPRES programme is described in detail in Section 3.3. In 1973/74, the Government of Indonesia (GOI) rolled out the massive SD INPRES programme during which the number of primary schools in the country nearly doubled. The SD INPRES programme was targeted at underserved areas that had relatively low primary enrollment prior to the programme. Exposure to the SD INPRES programme is determined by the intensity of school construction in district of birth at the time an individual started school.[45] An interaction term between a dummy variable indicating age in 1974 (year first SD INPRES schools completed) and programme intensity in district of birth (number of SD INPRES school built per 1000 children aged 5-14 in 1971) is used to capture SD INPRES programme exposure based on the 1995 intercensal survey of Indonesia (SUPAS). Only results for the sample of 31,061 male wage earners who were 23-45 years old in 1995 are discussed as these are most relevant to the findings in this chapter. The dependent variable of interest is monthly wages. In the first-stage regressions, SD INPRES programme exposure is significantly related to schooling and suggests that one additional school per 1,000 children increased schooling of exposed children by 0.2 years. The estimated return to schooling is nearly 7.0 percent treating schooling as exogenous but rises to 7.6-9.0 percent when instrumenting for schooling.

However, based on this variable for SD INPRES exposure, a district in which all SD INPRES schools were constructed in 1974, for example, is assigned the same programme intensity as a district in which same number of schools were constructed but not until 1978 (Hertz, Jayasundera 2007). A more recent study also uses the SD INPRES programme to assess the effect of the SD INPRES programme on schooling (and on intergenerational educational mobility) but based on a different measure of SD INPRES programme exposure (Hertz, Jayasundera 2007). The programme exposure variable in this case is the average number of SD INPRES schools constructed before each of the six years of primary schooling for an individual, assuming that she attended primary school between ages 7 and 12. This variable takes advantage of the fact that schools were constructed at different times in different districts thereby capturing additional variation in SD INPRES exposure. The data come from the Indonesia Family Life Survey (IFLS), the same data as those used in this chapter, and the effect of the SD INPRES programme on schooling attainment is examined separately for roughly 5,142 men and 5,742 women aged 24-50 years in 2000. Men exposed to the SD INPRES programme have significantly more schooling (0.30-0.56) whereas there is no effect of the programme for women. More parental schooling (mean of father's and mother's schooling) is significantly associated with more schooling for both men and women (0.50-0.63). An interaction for SD INPRES programme exposure and parental

[45] District of birth rather than district of schooling is used due to data limitations. However, it is district of schooling (which is highly correlated with district of birth) that determines programme exposure together with year of birth and programme intensity.

schooling is also included and suggests that the effect of SD INPRES programme exposure declines as parental schooling increases, i.e., children of parents with relatively less schooling seemingly benefitted more from the programme.

A recent study for Indonesia examines not only the average but also the marginal returns to upper secondary schooling (Carneiro, Lokshin et al. 2011). Distance from community of residence to nearest secondary school where distance takes the value zero if there is a school in the community of residence and one otherwise is used as an instrument for schooling together with interactions of distance with parental schooling, religion, age, and distance to a health center. It may be that individuals choose to locate close to a secondary school for reasons correlated with wages in which case distance to nearest secondary school is not a valid instrument. To counter this possibility detailed controls for individual and regional characteristics are included after which distance to secondary school is treated as exogenous. To measure schooling two variables are used. First, a dummy variable that is zero for less than upper secondary schooling and one for upper secondary schooling or more, and second, a continuous variable years of schooling, and results for both are reviewed. The sample comprises 2,608 men who were 25-60 years in 2000 who are wage-earners and the data come from the IFLS. In both sets of first-stage regressions, distance to secondary school is negatively and signficantly associated with schooling (results are not reported for the interaction terms) and father's and mother's schooling are positively and significantly related to schooling. For the regressions using the dichotomous secondary schooling variable the instrument is relevant based on the F-test of the significance of excluded instruments. But in the first-stage regressions using the continuous schooling variable the F-statistic is 3.6, which is relatively low with potential implications for bias (Bound, Jaeger et al. 1995, Staiger, Stock 1997). In the wage equations the dependent variable is log hourly wages. Using the dichotomous schooling variable the annualized ordinary least squares estimate of the return to upper secondary schooling is 9.0 percent compared to 12.9 percent for the instrumental variable estimate. When using years of schooling instead, the estimated return to upper secondary schooling is 9.6 percent not accounting for the endogeneity of schooling and the instrumental variable estimate is again larger at 15.7 percent. For the ordinary least squares estimates father's schooling is positively and significantly related to wages but based on the instrumental variable estimates it is not, which is similar to the studies above. This is discussed further in Section 3.6. The return to schooling may be heterogenous and the estimates suggest that this may be the case;

the estimated return to schooling is 26.9 percent for the average student in upper secondary school compared to 14.2 percent for the marginal student.

3.3 The Primary School System and SD INPRES Programme

This section first provides an overview of the primary school system and schooling in Indonesia prior to the SD INPRES programme. It then describes the design and implementation of the programme.

3.3.1 The Primary School System and Schooling in the 1970s

In the early 1970s before the massive primary SD INPRES school construction programme began, Indonesia had a population of 104 million living in 26 provinces on approximately 3,000 islands speaking roughly 250 different languages (Beeby 1979). The school system consisted of six years of primary school; three years of junior secondary school; and three years of senior secondary school. The legal minimum school starting age was six, but the majority of children began school at age seven and some at an even older age, primarily due to shortages of school places. The typical primary school had one principal and five to six teachers and between 150-300 students (Beeby 1979).

Indonesia had a relatively high capacity to enroll primary school aged children in 1971, with a gross primary enrollment ratio of 80 percent. However, due to large numbers of over-age students, only 68 percent of eligible children (net primary enrollment) were enrolled and rural children and girls were much less likely to be enrolled than urban students and boys respectively (World Bank 1975). Moreover, dropout rates were relatively high. 13 percent of children enrolled in primary school dropped out before grade three; 24 percent before grade four; and more than 55 percent before grade six. Out of those reaching grade six, only 61 percent graduated (Beeby 1979).

The majority of parents interviewed in education surveys at the time cited lack of money as the main reason for their child dropping out. The cost of sending a child to school including school fees imposed a barrier to entry for children from poor households, and when parents sent their children to school, they were frequently unable to keep them in school (Daroesman 1972).

There were two types of primary school fees in the early 1970s, entrance fees and monthly fees collected by parents' association for public schools. To increase transparency and accountability and reduce fees for the poor, a fixed, single parental contribution (SPP) was introduced in 1971 to replace the old fee system (Beeby 1979). The SPP comprised 1-2 percent of household income depending on costs in the area in which the household resided (World Bank 1975). In addition, there were other direct costs for uniforms, school supplies, and transport, and the opportunity cost of sending children to school.

As well as the high cost of schooling many parents considered three to four years of schooling adequate feeling that the literacy and numeracy skills achieved at these schooling levels were sufficient for the needs of their children (Beeby 1979).

The constraints of the primary school system were evident in schooling outputs and outcomes. In 1970, average years of schooling of the adult population (15 years and older) was only 2.8 years; 41 percent of the population had no schooling at all; and just 19 percent had completed primary schooling (World Bank 1975, Barro, Lee 2010). Merely 60 percent of the adult population was literate.

There were also large differences by location and gender: 79 percent of the urban population was literate but only 55 percent of the rural population and merely 49 percent of women compared to 71 percent of men. Moreover, among children whose fathers had professional occupations only 1 percent had no schooling, the corresponding proportions for children of fathers with mid-level occupations was 15 percent and for children of fathers with low-level occupations a staggering 84 percent (Table 28). Together these data underline the differences in schooling opportunities in Indonesia depending on gender, location, and socioeconomic group in the early 1970s.

3.3.2 The SD INPRES programme

The number of primary schools in Indonesia increased gradually over the period 1956-1965, but the construction effort slowed down after 1966 (UNESCO 1976). Indonesia's first national five-year development plan 1968/69-1973/74 (Repelita I) focused on economic development and its education component was concerned almost exclusively with vocational and technical training, while ignoring the overall structure of education including primary education (Beeby 1979, World Bank 1989). In 1971, two years before the start of the SD INPRES programme, there were approximately 61,000 public primary schools under the jurisdiction of the Department of Education in the country, and in addition, there were about 20,000 Madrasah primary schools (World Bank

1975).[46] Out of school-age children, 68 percent were enrolled, and the number of children nearing school age was increasing rapidly (Beeby 1979).

The hike in oil prices in 1973/74 expanded the size of the Indonesian government budget 2.5 times between 1973 and 1975, which substantially increased the Government of Indonesia's (GOI) ability to finance the extensive education reforms that were part of the second national five-year development plan (Repelita II) (World Bank 1989).[47]

An important aspect of the SD INPRES programme was that the new primary schools were to supplement rather than replace existing primary schools. The main objective was to provide school places for children who previously had no school to attend (Aziz 1990, Beeby 1979). Thus, the number of new schools to be built in each district was proportional to the number of children of appropriate school age not enrolled in 1971/72 (Chernichovsky, Meesook 1985, Aziz 1990).

The location of schools was chosen by the district chief executive (bupati) after consultation with the heads of sub-districts and primary school inspectors based on three criteria. Firstly, schools were to be located where children were currently unable to start first grade in existing schools, and preference should be given to low-income urban and remote areas (UNESCO 1976).

Each new school had three classrooms, furniture, equipment, and toilets and textbooks were provided (UNESCO 1976). To ensure that the quality of schooling did not suffer due to teacher shortages as primary enrollment increased the GOI recruited and trained a large number of teachers. At the time, 25,000 new teachers entered the labor market annually and there were thousands of teachers not working as teachers due to the hiring freeze imposed in 1968 (World Bank 1989). The freeze was abolished in 1973, the year the SD INPRES programme started, enabling the hiring of 18,000 new teachers and 41,000 temporary and part-time positions to be made permanent. As a result, student-teacher ratios remained nearly constant between 1973/74 (31.5:1) and 1978/79 (31.8:1) (Beeby 1979, World Bank 1989). Due to careful planning the quality of primary schooling did not seem to deteriorate during the SD INPRES programme implying that any effect of the programme operated through changes in quantity not quality of schooling (Beeby 1979, Duflo 2001, World Bank 1989).

---

[46] Most of these schools had a status similar to that of the schools under the Department of Education and followed the state curriculum.

[47] Under Repelita II Rp436 billion were allocated to education compared to Rp36.6 billion under Repelita I (World Bank 1989).

Table 29 shows the number of primary schools to be constructed as part of the SD INPRES programme: 62,000 schools or equivalently 2 schools per 1,000 children aged 5-14 in 1971. This constituted more than a doubling of the existing number of primary schools under the jurisdiction of the Department of Education in just five years.[48] The expansion of the school supply was staggered with the number of schools constructed gradually increasing over the course of the SD INPRES programme.

From 1960 until the start of the SD INPRES programme in 1973/74, average annual growth in primary enrollment was only 0.9 percent per year (World Bank 1989). However, during course of the SD INPRES programme the number of students enrolled in primary school increased dramatically from 13 million in 1972/73 to 19.1 million in 1978/79 (Figure 14).

Primary school fees were abolished for grades 1-3 in 1977 and for grades 4-6 in 1978 (Chernichovsky, Meesook 1985). The removal of official school fees may also have contributed to the rise in enrollment after 1977/78 but as Figure 14 shows, the rise in primary enrollment was steady before and for about one year after fee abolition when it slowed down.

Feasibility of the SD INPRES Programme as a Natural Experiment

To assess the feasibility of using the SD INPRES programme to capture exogenous variation in schooling a difference-in-difference (DID) estimator is used (Duflo 2001). The estimate of the effect of the SD INPRES programme on schooling is obtained as the difference in differences between individuals exposed and not exposed to the SD INPRES programme in high and low programme intensity regions respectively. To determine whether a district is assigned as high or low programme intensity a regression of the number of SD INPRES schools on the number of school-aged children for all districts are run (results not reported). All SD INPRES schools constructed had three classrooms, toilets, furniture, and equipment (see above). For positive residuals a district is defined as high programme intensity and for negative residuals as low programme intensity. In low programme intensity districts the mean number of SD INPRES schools constructed over the period 1973/74-1978/79 per 1,000 children aged 5-14 was 1.5 compared to 2.6 in high programme intensity districts.

[48] The number of schools planned corresponded very closely to the number of schools constructed (Beeby 1979, World Bank 1989).

As Table 30 shows, individuals in low programme intensity districts had more schooling (1.7 and 1.3 additional years respectively) on average than those in high programme intensity districts as expected since the SD INPRES programme targeted areas with low primary enrollment (Aziz 1990). Mean schooling attainment is higher also for exposed individuals (2.3 and 2.0 additional years respectively) than those not exposed in both high and low programme intensity districts. The DID estimate of the SD INPRES programme effect is 0.34 years of schooling. This is larger than a previous corresponding estimate for Indonesia of 0.12 years of schooling (Duflo 2001). The smaller estimate is from a different sample of men only whereas the sample in this chapter includes both men and women, and it may be that the average return to schooling is higher for women.

3.4 Sample Characteristics and Data

This section describes the sample, data, and instrumental variables used to examine the effect of the increase in primary school supply under the SD INPRES programme on schooling attainment and to assess the causal effect of schooling on wages in Indonesia. Variable descriptions and summary statistics for all variables are provided in Table 31 and Table 32.

3.4.1 Sample

The data used for the analysis come from the Indonesia Family Life Survey (IFLS), a repeated socioeconomic and health survey run by the RAND Corporation and the Center for Population and Policy Studies of the University of Gadjah Mada in Indonesia. The first three rounds of the survey: IFLS1, IFLS2, and IFLS3 conducted in 1993, 1997 and 2000 respectively are used. The 10,435 households interviewed in IFLS3 live in 13 of Indonesia’s 26 provinces, and represent approximately 83 percent of the Indonesian population in 1993 (Strauss, Beegle et al. 2004). The 13 provinces covered by the IFLS are: Sumatra Utara, Sumatra Barat, Sumatra Selantan, Lampung, DKI Jakarta, Java Barat, Java Tengah, DI Yogyakarta, Java Timur, Bali, Nusa Tenggara Barat, Kalimantan Selantan, and Sulawesi Selantan, which are the most densely populated areas in Indonesia (Annex figure 1).

The unit of analysis is men and women aged 25-50 years in 2000. This age group was chosen to include individuals not exposed and exposed to the SD INPRES programme, which started in 1973/74 and is used as an instrumental variable, and to ensure they were old enough to be working at the at the time of the survey.

Keeping observations for which data on wages (and all other variables used in the analysis) are available the sample is reduced from 10,320 to 3,146 (4,023) observations. The exclusion of self-employed workers is due to the lack of earnings data for this group. Means across wage and non-wage earners are similar except for the share of women, which is significantly larger among non-wage earners (34 percent) than among wage earners (56 percent), and that non-wage earners have significantly less schooling (1.5 years) than wage earners. Consequently, the findings may not apply to non-wage earners who may be systematically different from wage earners. A previous study of Indonesia finds some evidence of sample selection for with the SD INPRES programme impact being somewhat smaller for the full sample compared to the subsample of wage earners (Duflo 2001).

The data on SD INPRES programme school construction, the INPRES water supply and sanitation programme, and the primary school enrollment rate in each district in 1971 come from administrative records for 1973-1979 (BAPPENAS 1973, 1974, 1975, 1976, 1978). The IFLS data were merged with the administrative data, which was very time consuming as several of the 283 districts (kabupaten) changed name, some more than once, between 1973/74 when the SD INPRES programme started and 2000 when the third IFLS wave was fielded.

3.4.2 Data

Schooling

Schooling attainment is the dependent variable in the first set of regressions that examine the effect of the SD INPRES programme on schooling in general and for poor children and girls in particular. It is also the main independent variable of interest in the assessment of the effect of schooling on wages. In this chapter, schooling is defined as years of formal schooling and is constructed by combining the reported highest grade and level of formal schooling completed.

In the sample, mean years of schooling is 8.6 with a standard deviation of 4.5 years. Figure 15 shows the heaping of schooling around the completion of each schooling level. Almost 6 percent have no schooling; 20 percent have completed primary schooling; 28 percent have completed junior secondary school; 12 percent have completed senior secondary school, and 12 percent have a university education.

Examining schooling for those exposed and not exposed to the SD INPRES programme reveals differences across the two groups. Mean schooling is higher for those exposed to the programme, 9.5 years compared to 7.4 years for those not exposed (Table 33). The difference in schooling across the exposed and unexposed groups is significantly different indicating that the SD INPRES programme, without controlling for any other factors that also influence schooling, is associated with more schooling in the sample.

Schooling is arguably endogenous in the analysis of wages and estimates of the effect of schooling on wages obtained by ordinary least squares estimation will therefore be biased and inconsistent (Card 1999, Card 2001). There are two main sources of bias in this context. The first is unobserved heterogeneity such as differences in ability or family background that influence both schooling and wages and for which the direction of bias is ambiguous (Griliches 1977, Sander 1992). Second, schooling may be measured with error, especially as recall periods become longer, and due to grade repetition, which tend to bias the estimates of the effect of schooling on wages downward (Behrman, Deolalikar 1991, Griliches 1977, McCrary, Royer 2011).

To obtain consistent estimates of the effect of schooling on wages this chapter uses SD INPRES programme exposure and an interaction term for SD INPRES programme exposure and father's schooling to capture exogenous variation in schooling.

Wages

The dependent variable in wage equation is the natural log of monthly wages measured in local currency. In addition to monthly wages, the number of hours normally worked per week (and last week) and the number of weeks normally worked per year, which could be used to construct hourly wages, are reported. However, dividing monthly wages by the number of hours typically worked each week times the number of weeks typically worked is likely to increase measurement error. Therefore, monthly wages is the preferred variable, although it to some extent confounds the effects of labour supply and human capital. Individuals with more schooling generally work more hours and as a result, the estimates of the return to schooling tend to be higher for annual wages than hourly wages (Card 1999). To address the issue of part-time workers, about 28 percent of the sample, a dummy variable that captures part-time work is included in the model.

Wages are roughly normally distributed with a mean of Rupiah 220,000 per month for the sample. To put this into some context, mean wages for Indonesian wage earners and self-employed workers in 1999 were roughly Rupiah 347,000 per month (self-employed workers are excluded from the sample as discussed earlier) and the Indonesian poverty line, was Rupiah 16,700 per month in 1990 (Dhanani, Islam 2004).

SD INPRES programme exposure

The intensity and timing of the construction of SD INPRES schools varied substantially across districts. In this chapter, an individual's programme exposure is determined by the intensity of SD INPRES school construction and age at the time SD INPRES schools were constructed in her district following the approach in Duflo (2001). However, the exposure measure used here defines programme intensity differently and uses more of the variation in the timing of school construction to attempt to capture more of the exogenous variation in SD INPRES school construction for each individual (Figure 16).

Duflo (2001) was first to use the SD INPRES programme to evalute the effect of increased school supply on schooling and the return to schooling for men in Indonesia as discussed in Section 3.2. Exposure to the SD INPRES programme was constructed as an interaction term between a dummy variable indicating age in 1974, the year the first SD INPRES schools were completed, and programme intensity defined as the number of SD INPRES school built per 1000 children aged 5-14 in 1971 (Figure 16). Individuals 12-14 years old in 1974 constitute the unexposed group and individuals who were 2-6 years old in 1974 comprise the exposed group (those who were 7-11 years old in 1974 are excluded from the sample). However, based on this exposure variable an individual in a district in which all SD INPRES schools were constructed in 1974, for instance, is assigned same programme exposure as an individual in a district in which same number of schools were constructed but not until 1979 (Hertz, Jayasundera 2007). The SD INPRES programme was also used to examine the effect of increased school supply on schooling attainment for men and women (but not wages) based on IFLS data, also discussed in Section 3.2 (Hertz, Jayasundera 2007). This study treats individuals born in 1950-61 as not exposed to the SD INPRES programme, those born 1962-71 as partially exposed and those born 1972-76 as fully exposed (Figure 16). The programme intensity in this case is the average number of SD INPRES schools constructed before each of the six years of primary schooling for an individual. Based on this measure of SD INPRES programme exposure, in a district that constructs three schools each year for three years, exposure would be assigned as three schools for an individual regardless of whether she is of school-starting age in year 1, year 2, or year 3. In another district, which also constructs nine schools over three years but builds all nine schools in year 3, an individual would be assigned exposure of 3 schools regardless of whether she starts school in year 1, 2 or 3 although only the individual who starts school in year 3 is exposed to the SD INPRES programme.

The variation in SD INPRES school construction intensity by district and over time also informs the measure of programme exposure used in this chapter (Figure 16). Individuals aged 7 years or younger when the first SD INPRES school was built in their district comprise the exposed group and those aged 13 years or older form the not exposed group (individuals aged 8-12 years (over-age) when SD INPRES schools were constructed are excluded from the main sample).[49] Because the number of SD INPRES schools constructed by district was an increasing function of time the measure of programme intensity is the cumulative number of SD INPRES schools constructed per 1,000 children aged 5-14 years in a given district in the year before an individual started primary school. This programme intensity measure uses the fact that individuals born later in a given district, conditional on them being of appropriate age when the schools were constructed, have higher SD INPRES programme exposure and avoids the potential problem associated with using the average number of schools discussed above. Moreover, due to difficulties of assigning correct programme exposure to over-age school starters (see below) instead of assuming that all individuals born 1962-71 are partially exposed as in Hertz and Jayasundera (2007), some of these individuals may have been fully exposed, some partially and some not at all, this chapter excludes potentially partially exposed individuals from the main sample.

An advantage of the IFLS data is that individuals who moved from their district of birth between birth and age 12 can be identified. Among all individuals born between 1950 and 1975, 13 percent moved before turning 12 years old, the age at which primary school is completed, and these observations are excluded from the sample to ensure correct assignment of SD INPRES programme exposure.[50] The group means for the movers and non-movers are similar except that the share of women is significantly

---

[49] The regressions are also run for another sample that includes over-age school starters and assigns them full programme exposure (Sample 2) as a robustness check.
[50] Assuming that children started school at the official school-starting age and that there is no grade repetition.

larger and mean schooling significiantly lower among the movers, which implies that movers and non-movers may be systematically different. This is similar to findings for rural Philippines (Maluccio 1998). Consequently, if movers are systematically different from non-movers, for example, poorer people may be more likely to move and have less schooling, the findings will not apply to the former.

The issue of children who were over-age, i.e., older than seven when starting school, needs to be considered when constructing the SD INPRES programme exposure variable. The IFLS contain data on the age at which an individual started school, which could potentially be used to determine programme exposure for over-age individuals. However, among over-age school-starters there is substantial variation in the years of schooling completed, with a large proportion not completing the full six years of primary school, and a large fraction only obtaining three years of schooling or less. This makes it difficult to assign over-age school-starters correct programme exposure. Therefore, individuals who were 8-12 years old in the year in which SD INPRES schools were constructed in their district of schooling are excluded. Consequently, individuals unexposed to the SD INPRES programme are those 13 years or older in the year SD INPRES schools were built in their district, and exposed individuals are those seven years or younger at the time of school construction (sample 1).

However, as a robustness check, a sample (sample 2) that includes individuals who were over-age at the time SD INPRES schools were constructed in their district is also used and these individuals are treated as fully exposed to the SD INPRES programme.

That schooling is likely endogenous in the analysis of schooling and wages was discussed above. For the SD INPRES programme to be a valid instrument it must be highly correlated with the endogenous variable schooling (relevance) and it must not be correlated with the error term in the wage equation (exogeneity). Because of the massive size of the SD INPRES programme and the rise in primary enrollment simultaneous with its rollout, it is reasonable to assume that the programme affected schooling attainment and therefore constitutes a relevant instrument. Instrument relevance is formally tested in Section 3.6.

For the SD INPRES programme to be exogenous it must only affect wages through schooling. Given that the main determinant of the allocation of SD INPRES schools to each district was the number of school-age children not enrolled in primary school before the programme started, the estimates may be biased due to correlation between prior enrollment rates and the programme. To control for the school allocation rule an interaction term between year of birth and the proportion of children enrolled in 1971 in each district is included in the model. In addition, the number of school-age children in 1971 is negatively related to the number of SD INPRES schools constructed per child. The likely reason is that, on average, districts with more school-age children were those with more schools before the programme started (less remote areas) and therefore these districts were allocated fewer SD INPRES schools since by construction the programme targeted underserved areas. To prevent this correlation from biasing the estimates downward an interaction term between a cohort of birth dummy and the proportion of the population aged 5-14 in 1971 is used (Duflo 2001).

The identifying assumption is that the observed increase in schooling would not have varied systematically across districts in the absence of the SD INPRES programme, which cannot be tested directly. However, a falsification exercise examining the difference between SD INPRES and non-SD INPRES districts in schooling attainment of two groups (individuals aged 12-17 in 1974 and individuals aged 18-24 in 1974) neither of which was exposed to the SD INPRES programme finds no signficant difference (Duflo 2001).

The identification assumption also fails to hold if the allocation of other contemporaneous public programmes that affected schooling is correlated with the allocation of SD INPRES schools. This may be the case for the Repelita II water supply and sanitation programme rolled out concurrently with the SD INPRES programme. If allocation of the SD INPRES programme and the water supply and sanitation programme are correlated, improved access to clean water and sanitation may lead to healthier children who obtain more schooling, affecting schooling in the same direction as the SD INPRES programme. Alternatively, the water supply and sanitation programme may act as a proxy for poor child health if it was allocated to districts with high morbidity and mortality (as was the case in Indonesia), and children of poorer health obtain less schooling, thereby affecting schooling in the opposite direction to the SD INPRES programme. To control for this possibility, interaction terms for year of birth and the allocation of the water supply and sanitation programme are included (Duflo 2001).

The high cost of schooling discussed above led the GOI to abolish school fees to ensure that school-age children would not be prevented from enrolling in the new SD INPRES schools due to cost. To make sure that fee abolition does not bias the estimates upwards a dummy variable for individuals whose schooling decisions were potentially affected by fee abolition, those who started school after 1977/78, is included, which previous studies have not done.

To account for district-specific effects that may affect wages the model includes dummy variables for each district. Likewise, to control for year-specific effects dummies for year of birth are used. After including the controls discussed above it is argued that the exogeneity assumption is valid. The validity of the instrumental variables is also formally tested in Section 3.6.

The average number of SD INPRES schools constructed was two schools per 1,000 children over the period 1973/74-1978/79. In the sample, mean SD INPRES programme exposure is 0.64 (Table 32). Mean SD INPRES programme exposure and schooling for individuals exposed and not exposed respectively to the SD INPRES programme are shown in Table 33.[51] For exposed individuals mean programme exposure is 1.1 schools per 1,000 children (the largest cumulative number of SD INPRES schools built per 1,000 children in any district before an individual started school is 6.4) and on average, they have 2.1 more years of schooling than those not exposed, and these differences are statistically significant.[52]

Father's Schooling

Socioeconomic background influences schooling attainment and wages (Dearden 1999, Freeman 1986, Schultz 1988,). Socioeconomic background, which is in this chapter is captured by father's schooling, is included to help account for parental investments in their children's human capital that are not fully captured by schooling and for unobservable ability (Strauss, Thomas 1995). To examine whether the SD INPRES programme was successful in targeting children from less advantageous socioeconomic backgrounds and whether increased schooling translates into higher wages regardless of socioeconomic background an interaction term between father's schooling and the programme is included.[53]

Father's schooling is chosen rather than household income or consumption for two reasons. First, no data on household income or consumption for the period 1950-75 are

---

[51] By construction, programme exposure of the unexposed group is zero. The reason for the difference in the mean number of schools built per 1,000 children across districts and mean SD INPRES programme exposure is that the latter uses the cumulative number of schools constructed in the year before starting school per 1000 children as the numerator and the number of children in that year as the denominator whereas the former uses the period total (1973/74-1978/79) number of schools as the numerator and the number of children in 1971 as the denominator.

[52] The highest number of SD INPRES schools built per 1,000 children in any district is 6.4.

[53] Some studies use parental schooling to control for ability (Card 1999).

available for individuals exposed and not exposed to the SD INPRES programme. Second, it tends to constitute a broader and more stable measure over time of socioeconomic background than income or consumption.

Father's schooling is defined as formal years of schooling attained and was obtained by linking individuals in IFLS with their parents. The mean schooling of fathers is 5.2 years with a standard deviation of 3.9 years (Table 32), which is lower than mean schooling of their sons (8.9 years).[54] Moreover, among the fathers, 20 percent have no schooling; 36 percent have completed primary schooling; and 10 percent and 9 percent have completed junior or senior secondary schooling respectively. The large difference in schooling attainment for fathers and sons is largely a result of the SD INPRES programme during which primary enrollment became nearly universal.

Gender

In developing countries, women tend to have less schooling than men and earn lower wages including in Indonesia (Behrman, Deolalikar 1995, Oey-Gardiner 1991, Strauss, Thomas 1995, UNESCO 2004). Reasons for gender differences in schooling and wages include the possibility that families invest more in the schooling of sons if the return to their schooling is higher than that for daughters, or if sons are expected to provide security in old age (Strauss, Thomas 1995). Another possibility is that the cost of sending girls to school is relatively higher than that for boys if girls help at home. However, the opposite may also be the case, that the cost of schooling for boys is relatively higher if they help with agricultural work for example (Oey-Gardiner 1991). Differences in schooling may also reflect cultural norms that value the schooling of boys and girls differently (Oey-Gardiner 1991, Strauss, Thomas 1995).

In the sample, 34 percent are women (Table 32). To capture any differences in schooling attainment and in wages by gender a dummy variable that takes the value one for women is included in the model. In addition, to examine whether the SD INPRES programme affected the schooling of men and women differentially, an interaction term between gender and SD INPRES programme exposure is used.

3.5 Conceptual Framework and Empirical Methodology

[54] Mean schooling of fathers and sons are compared rather than that of fathers and their sons and daughters as women tend to have less schooling than men in Indonesia.

This section outlines the human capital model that underlies the empirical model. It then discusses the first-stage schooling equation, the wage equation, and estimation methods.

3.5.1 Conceptual Framework

The empirical model of schooling and wages is based on a standard human capital model (Becker 1965, Becker 1964, Ben-Porath 1967, Freeman 1986, Mincer 1958). In the basic model below reproduced from Freeman (1986), individuals invest in human capital through time spent in school, and the return to schooling, r, is given by the continuous investment function:

$$(1) \quad \int_0^n W_t e^{-rt} = \int_s^{n_s} W_{st}\, e^{-rt} + \int_0^s (W_{st}^{'} - D) e^{-rt}$$

Here n is the number of years individuals with no schooling work and $n_s$the number of year individuals with schooling work. W_t is earnings for individuals without the relevant schooling. W_{st} is earnings for individuals with schooling after completing school and $W_{st}^{'}$ is earnings during their time in school. D is the direct cost of schooling (e.g., school fees, materials) and s is the years of schooling. In the model, the discounted present value of earnings of individuals who do not go to school is captured by the first term. The second term is the discounted present value of earnings for those who invest in schooling after they have completed schooling and the third term is their net earnings during their time in school. Under the assumption that net earnings while in school average out to zero; workers with schooling retire later; and earnings are not influenced by age or experience, equation 1 becomes:

$$(2) \quad \int_0^n W\, e^{-rt} = \int_0^{n+s} W_s e^{-rt}$$

or

$$(3) \quad \frac{W_s}{W} = \frac{e^{-r(n+s)} - e^{-rs}}{e^{-rn} - 1} = e^{rs}$$

Taking logs yields equation 4:

(4) $\log w_s = \log w + rs + (other\ terms)$

This is the standard semi-log earnings function used to estimate the schooling and earnings relationship and in which r captures the return to schooling.

3.5.2 Estimating the Effect of the SD INPRES Programme on Schooling

To examine whether exposure to the SD INPRES programme is associated with more schooling, and whether there is a differential impact of the programme depending on socioeconomic background and gender, the schooling equation is estimated by ordinary least squares. This equation also provides the first-stage for the instrumental variable estimation.

In the first-stage, schooling is regressed on all exogenous variables including SD INPRES programme exposure and the interaction of SD INPRES schooling and father's schooling to generate the fitted values for the second-stage wage regression. The equation also includes controls for father's schooling, gender, an interaction of gender with SD INPRES programme exposure, a dummy variable for fee abolition, and district and year of birth specific effects.

Interactions of year of birth with primary enrollment and with the proportion of children in each district in 1971 are included to control for the school allocation rule (Duflo 2001). Finally, an interaction between year of birth and the water supply and sanitation programme is used to control for the possibility that this public programme contemporaneous to the SD INPRES programme also affected wages (Duflo 2001). The first-stage schooling equation is given by:

(5) $$S_{ijk} = c_0 + \alpha_j + \beta_k + \delta_1 P_{ijk} + \delta_2 P_{ijk} E_i + \delta_3 E_i + \delta_4 W_i + \delta_5 P_{ijk} W_i + \delta_6 F_i + \delta_7 C_j + \varepsilon_{ijk}$$

i=individual
j=district of birth
k=year of birth
c_0 = constant
α_j = district of birth specific effect
β_k = year of birth specific effect
S_{ijk} = years of schooling
P_{ijk} = SD INPRES programme exposure: cumulative number of schools constructed per 1,000 children aged 5-14 in district of schooling in year prior to starting primary school
$P_{ijk}E_i$= interaction between SD INPRES programme exposure and father's schooling
E_i = completed years of schooling of father
W_i= dummy variable equal to one if woman
$P_{ijk}W_i$= interaction between SD INPRES programme exposure and gender

F_i= dummy variable for fee abolition
C_j = interactions between year of birth and primary enrollment rate; year of birth and proportion of school-age children; and year of birth and water supply and sanitation programme in 1971 in each district
ε_{ijk}=unobservables (e.g., innate ability)

3.5.3 Estimating the Return to Schooling

The wage equation is first estimated by ordinary least squares using schooling to generate benchmark results for the relationship between schooling and wages.[55] To assess if more schooling causes higher wages the fitted values of schooling from the first-stage are used in the second-stage wage equation:

$$(6) \quad w_{ijk} = c_0 + \alpha_j + \beta_k + \gamma_1 \hat{S}_{ij} + \gamma_2 E_i + \gamma_3 W_i + \gamma_4 P_{ijk} W_i + \gamma_5 F_i + \gamma_6 C_j + \upsilon_{ijk}$$

w_{ijk}= monthly wages (natural log)
$\hat{S}_{ij}$= predicted schooling
υ_{ijk}=unobservables
All other variables defined as above

The identification comes from $P_{ijk}$and $P_{ijk}E_i$ in the first-stage and γ_1 captures the causal effect of schooling on wages. This model estimates the total effect of schooling but does not reveal the channels through which schooling influences wages.

3.6 Results and Discussion

This section begins by presenting the first-stage schooling results and tests of instrument relevance and validity. These results are of key interest as they reveal the effect of the increase in school supply on schooling attainment. Next, the estimates of the effect of schooling on wages are discussed.

3.6.1 The Effect of School Supply on Schooling Attainment

The results from the first-stage schooling regressions are shown in Table 34. Columns 1-3 contain the estimates for the main sample (sample 1) and columns 3-5 the results for the sample that includes children who were overage (8-12 years) at the time SD INPRES schools were constructed in their district and treats them as fully exposed (sample 2).

---

[55] This wage equation (not shown) is the same as the second-stage equation shown above except years of schooling are used rather than the fitted values from the first-stage schooling equation.

The preferred estimates are those in columns 2 and 5 that control for fee abolition, part time work, and allocation of the water supply and sanitation programme in addition to the other covariates discussed above. The regressions are also run without controlling for part-time work and this does not change the results (columns 3 and 6).

The results in Table 34 show that individuals exposed to the SD INPRES programme have significantly more schooling (columns 1-3). A one standard deviation increase in exposure (0.82 schools per 1,000 children of school-age) is associated with 0.8 additional years of schooling. This effect is smaller when not controlling for the water supply and sanitation programme similar to the findings in a previous study (Duflo 2001). As discussed in Section 3.4, the water supply and sanitation programme may capture poor child health and targeted some of the same districts as the SD INPRES programme. If children of poorer health acquire less schooling, the water supply and sanitation programme will affect schooling in the opposite direction to the SD INPRES programme.

SD INPRES exposure is also associated with more schooling for the sample that includes overage school-starters (columns 4-6) but the effect is somewhat smaller (0.6 years), which would be consistent with the notion of some observations being assigned as fully exposed although they were only partly, or not exposed at all to the programme. The positive effect of the SD INPRES programme on schooling is similar to other evidence for Indonesian men (Duflo 2001, Hertz, Jayasundera 2007).

The interaction term for SD INPRES programme exposure and father's schooling is negatively and significantly related to schooling and of similar size across the two samples (-0.107). This provides some evidence that children from less favourable socioeconomic backgrounds, as measured by father's schooling, benefitted more from the SD INPRES programme. This finding is similar in size and direction to those of an earlier study of Indonesia and a study of Honduras for men (Bedi, Gaston 1999, Hertz, Jayasundera 2007). Moving on to the second interaction term between SD INPRES programme exposure and gender, it is positively and significantly associated with schooling (0.64 and 0.55 in columns 2 and 5 respectively) indicating that girls benefitted relatively more from the programme. Thus, it appears poorer children and girls benefitted relatively more from the SD INPRES programme. This comes as no surprise since the programme was targeted at underserved areas where initial enrollment was low, and because children from poorer families and girls were less likely to be enrolled prior to the SD INPRES programme as described in Section 3.3.

However, in general, being a girl is significantly associated with fewer years of schooling but this is partly offset by exposure to the SD INPRES programme (columns 2 and 5). Individuals whose fathers have more schooling have significantly more schooling and a one standard deviation (4 years) rise in father's schooling is significantly associated with roughly 2.2 additional years of schooling. This is similar to previous studies for Indonesia and other countries (Bedi, Gaston 1999, Carneiro, Lokshin et al. 2011, Hertz, Jayasundera 2007).

Fee abolition is controlled for in columns 2, 3, 5, and 6. However, there is no discernible independent effect of abolition of fees on schooling, which suggests that it did not increase schooling as discussed earlier (Figure 11).

The instrumental variables are relevant with the F-statistic for the excluded instruments ranging from 12.5 to 17.1 for the main sample and from 14.8 to 16.0 for the larger sample including overage school-starters (Bound, Jaeger et al. 1995, Staiger, Stock 1997). The instruments are also valid across all specifications and across the two samples based on the Hansen J-test (columns 4-6 in Table 35 and Table 36 respectively).

As a robustness check, wages are regressed on the SD INPRES programme exposure variable and the interaction of the programme and fathers' schooling for the main sample and the sample including overage school-starters. The coefficients on the excluded instruments are not significant, which further indicates that the SD INPRES programme affected wages through schooling and not directly.

School supply and employment and completion outcomes

An additional set of regressions that examine if the SD INPRES programme is correlated with other outcomes related to schooling are shown in Annex table 1-Annex table 8. The main sample is first divided by gender and then by socioeconomic background (SES) at the time an individual attended school. The latter is captured by father's level of schooling because no household income or asset data are available for the sample for the period during which the SD INPRES programme was rolled out (see above).[56] The first outcome variable is employment type (wage earner or self-employed) and the equations are estimated using probit models. The second outcome

[56] The sample is split into three groups: Low socioeconomic background (low SES) when fathers have no schooling, middle socioeconomic background (middle SES) when fathers have some or complete primary schooling and high socioeconomic background (high SES) when fathers have some secondary schooling or more.

variable is occupation which is divided into three broad categories: blue collar, services and white collar, and multinomial logit regressions are estimated with blue collar as the reference category.[57] The third outcome variable is primary school completion and the fourth is secondary school completion; these equations are estimated using probit models.

The findings suggest that men exposed to the SD INPRES programme are significantly more likely to be wage earners than self-employed but not women (Annex table 1). The SD INPRES programme also appears to reduce the relative advantage in terms of being in wage employment (as opposed to being self-employed) for men with more educated fathers. Both men and women whose fathers have more schooling are significantly more likely to be wage earners.

When it comes to occupation, for women there is no significant relationship between SD INPRES exposure and occupation but men who were exposed to the programme have a higher relative risk of being in a white collar than a blue collar occupation holding other factors constant (Annex table 2). The more schooling an individual's father has, the larger the relative risk that the individual has either a services or white collar occupation than blue collar occupation is. This effect is however reduced to some extent for both men and women exposed to the programme and whose fathers have more schooling similar to the finding on wage employment above (Annex table 1). Together this suggests that the SD INPRES programme raised intergenerational mobility in terms of occupation and employment type. In a similar vein, earlier evidence for Indonesia finds that the SD INPRES programme increased intergenerational education mobility for men (Hertz, Jayasundera 2007).

There is no significant effect of SD INPRES programme exposure on primary completion for men nor for women (Annex table 3) but individuals whose fathers have more schooling are more likely to complete primary school in line with expectations. That there is no relationship between SD INPRES programme exposure and primary completion suggests that the programme may have worked primarily by increasing enrolment (including of poorer students, see above) but did not necessarily improve retention. This would be consistent with the SD INPRES programme being a supply-side intervention that among other things reduced the distance to school but left

[57] The blue collar category includes agriculture, mining, construction, manufacturing and utilities, the services category comprise services, communications, and transportation, and the white collar category consists of finance, business and social services (includes teachers and civil servants).

demand-side constraints such as the cost of schooling largely unchanged. For secondary school completion, there is evidence that men exposed to the programme are more likely to complete secondary school, and similar to primary completion, individuals whose fathers are more educated are more likely to complete secondary school (Annex table 4).

Dividing the sample by socioeconomic background, men in the high SES group exposed to the SD INPRES programme are significantly more likely to be wage earners (Annex table 5). Women in all SES groups are significantly less likely to have wage employment, and for women in the middle SES group exposed to the SD INPRES programme this probability is further reduced.

In terms of occupation, men from the low SES sample who were exposed to the SD INPRES programme have a significantly higher relative risk of being in a white collar occupation than a blue collar occupation whereas women in the low SES exposed to the SD INPRES programme have a significantly lower risk of having a white collar job than blue collar job. There is no significant association between SD INPRES exposure and occupation for the middle and high SES samples (Annex table 6) but women in the middle and high SES sample have significantly higher relative risks of having a white collar than blue collar occupation.

When dividing the sample by SES, the primary and secondary school completion models can only be estimated for the low and middle SES samples as there is too little variation in these outcome variables for the high SES group (for instance, only 3.5 percent of the sample have not completed primary school). For the low and middle SES samples, SD INPRES exposure is not significantly correlated with primary school completion (Annex table 7) and women are significantly less likely to complete primary school than men. However, SD INPRES programme exposure partly offsets this negative gender effect for women in the low SES sample.

In terms of secondary school completion women in the low SES group are significantly less likely to complete secondary school but this is not compensated by SD INPRES exposure as was the case for primary completion. These findings suggest that the programme affected schooling of the poorest women mainly at primary level. For secondary school completion for the middle SES, individuals exposed to the SD INPRES programme are significantly more likely to complete than those not exposed but again women are significantly less likely to complete than men (Annex table 8).

3.6.2 Schooling and Wages: Ordinary Least Squares and Instrumental Variable Estimates

The results for the wage equation estimated by ordinary least squares and instrumental variables respectively are shown in Table 35 for the main sample and in Table 36 for the sample, which includes overage school-starters. In each table, the preferred specifications are columns 2 and 5, and these are discussed below.

Schooling causes higher wages and an additional year of schooling raises monthly wages by 13.5 percent in the main sample and by 13.2 percent in the sample with overage school-starters. These estimates are larger than existing evidence for Indonesia, which finds returns of about 9.0 percent for men also using monthly wages (Duflo 2001). The ordinary least squares estimates are smaller: one more year of schooling is significantly associated with wages that are 10.4 percent and 10.3 percent higher for the two samples. Whether the dummy variable for individuals who work part-time is included or not does not affect the results. Possible reasons for the difference in ordinary least squares and instrumental variable estimates are discussed below.

Interestingly, father's schooling is not related to wages in the instrumental variable regressions but it is in the ordinary least squares regressions where one additional year of schooling for fathers is associated with significantly higher wages (2.2 percent). This is similar to existing evidence that also finds that father's schooling is significantly associated with wages when assuming that schooling is exogenous but not when treating it as endogenous (Bedi, Gaston 1999, Carneiro, Lokshin et al. 2011). This result suggests that the effect of father's schooling may be operating mainly through formal investment in children's schooling and have no independent effect on wages (i.e., father's schooling does not capture ability).

As expected women earn significantly less than men, their wages are between 5.4 percent and 5.7 percent lower on average. An earlier study also finds gender differences in earnings and wages in Indonesia (Behrman, Deolalikar 1995). The interaction of SD INPRES programme exposure and gender is only significant when part-time work is not controlled for (column 3) in the main sample and using ordinary least squares. Still, the SD INPRES programme raises schooling by approximately 0.64 additional years for exposed women. Combining this schooling effect of the SD INPRES programme with the average effect of increased schooling on wages (0.135) implies that the SD INPRES programme reduced the average gender wage gap by roughly 9 percent.

Recalling the earlier findings, men exposed to the SD INPRES programme have a significantly higher probability of completing secondary school and of having wage employment and also have a significantly higher relative risk of being in a white collar occupation (Annex table 1, Annex table 2, Annex table 4). Men in the high SES group are also significantly more likely to be in wage employment and men in the low SES group to have a white collar job (Annex table 5, Annex table 6). This implies that for men higher wages caused by more schooling may come about partly by raising secondary completion and by placing men in wage employment and white collar jobs.

Studies of the impact of schooling on wages that use instrumental variables, particularly supply-side interventions, frequently find that the instrumental variable estimates are larger than the ordinary least squares estimates (Ashenfelter, Harmon et al. 1999, Card 2001). There are two main reasons for the difference in the estimates. First, measurement error in schooling may induce a downward bias in the ordinary least squares estimate (Griliches 1977). Second, returns to schooling may be heterogeneous (Imbens, Angrist 1994).

In the Indonesian case, those who decided to attend primary school in response to the increase in the primary school supply may have relatively higher returns to schooling than those whose schooling decisions were not affected by the programme in which case the instrumental variable estimator gives the local average treatment effect. Nevertheless, since the SD INPRES programme affected such a large proportion of the sample the local average treatment effect should approach the average treatment effect (Devereux, Fan 2011). For Indonesia, the larger estimate of return to schooling obtained when accounting for schooling being endogenous is probably the result of a combination of measurement error and heterogeneous returns (Maluccio 1998, Carneiro, Lokshin et al. 2011).

3.7 Conclusion

This chapter examines the effect of increases in school supply on schooling attainment and whether more schooling causes higher wages for men and women in Indonesia. The findings suggest that the SD INPRES programme significantly increased schooling both for men and for women, and that women benefitted more than men as did individuals from less advantageous socioeconomic backgrounds. Thus, it appears that the SD INPRES programme contributed to a narrowing of schooling gaps by gender and by socioeconomic background in Indonesia. This suggests that when there is unmet demand for schooling supply-side school infrastructure policies can be successful in raising the schooling of traditionally marginalized groups. Moreover, the SD INPRES programme appears to mainly have affected the schooling of women at the primary level and for men at the secondary level.

The chapter also finds that additional schooling causes higher wages both for men and for women but that there may have been an added positive effect on wages for women through the additional schooling induced by the SD INPRES programme.

Further, men have a significantly higher relative risk of having a white collar than blue collar occupation and a greater probability of being in wage employment rather than self-employed which suggests that more schooling partly causes higher wages through occupation choice and employment type.

The SD INPRES programme also appears to raise intergenerational mobility: It reduces the relative advantage in terms of having a wage earning job and not having a blue collar job for those with more educated fathers. Moreover, women are generally less likely than men to complete primary school but for the poorest women this negative gender effect is partly offset by SD INPRES programme exposure.

Moreover, the findings, in line with existing evidence, underline the importance of accounting for schooling endogeneity when assessing the effect of schooling on wages.

Two important issues are not dealt with in this chapter. First, the results presented in this chapter are for wage earners only, and the returns to schooling may differ for non-wage earners and currently no evidence on this group exists for Indonesia. Second, it is possible that the returns to schooling are heterogeneous in Indonesia, which is an issue that deserves more attention.

Table 28. Literacy and schooling by location, gender, and father's occupation, 1970-71

	Percent of population		
	literate	no schooling	complete primary schooling
total	59.6	41	19.4
by gender			
female	49	51.6	15.4
male	70.8	29.8	23.6
by location			
rural	55.2	45.2	27.4
urban	79.1	22.3	17.7
by father's occupation[1]			
high	-	1.0	-
medium	-	14.9	-
low	-	84.2	-

Note: 1. These data are from a small, non-representative sample so are indicative only.

Source: (World Bank 1975)

Table 29. SD INPRES programme investments, 1973/74-1978/79

	SD INPRES programme investments			
	New primary schools	Primary school books (millions)	SD INPRES allocation (billions of rupiah)	SD INPRES allocation as % of total public spending
1973/74	6,000	6.6	17.2	1.5
1974/75	6,000	6.9	19.7	1.0
1975/76	10,000	7.3	49.9	1.9
1976/77	10,000	8.6	57.3	1.6
1977/78	15,000	7.3	85.0	2.0
1978/79	15,000	8.5	111.8	2.1
Total	62,000	45.2	341	-

Note: All schools constructed had 3-classroom blocks. Textbooks in mathematics and Bahasa Indonesia.

Sources: (World Bank 1989, World Bank 1984).

Table 30. Mean schooling by SD INPRES programme intensity and exposure for sample

	Schooling		
SD INPRES programme	not exposed	exposed	difference
high intensity	6.4	8.7	-2.3*
low intensity	8.1	10.0	-2.0*
difference	-1.7*	-1.3*	-0.34*

Note: *statistically different at one percent.

Low/high programme intensity determined by a regression of the number of schools constructed by district on the number of school age children (5-14 years) in that district. For positive residuals a district is defined as high intensity and for negative residuals as low intensity.

Program exposure defined as cumulative number of SD INPRES schools constructed per 1,000 children aged 5-14 in an individuals' district of birth in year before started primary school.

Table 31. Variable description

Variable	Description
district	Dummy variable for respondent's district of birth
father's schooling	Father's years of completed schooling
fee abolition	Dummy variable if started school after primary school fee was abolished
part-time	Dummy variable for respondents who work part-time
proportion enrolled in 1971	Fraction of children enrolled in primary school in each district in 1971
proportion of children in 1971	Fraction of children aged 5-14 in each district in 1971
schooling	Years of completed schooling
SD INPRES pgm exposure	Exposure to SD INPRES school construction programme
wages	Self-reported monthly wages in local currency (natural log)
water supply and sanitation	Water supply and sanitation programme spending by district 1973-78
woman	Dummy variable equal to one for women
year of birth	Dummy variable for respondent's year of birth

Table 32. Summary statistics

Sample 1 (N=3146)	Mean	Std. Dev.	Min	Max
father's schooling	5.2	3.9	0	18
proportion enrolled in 1971	0.17	0.06	0.03	0.61
proportion of children in 1971	0.27	0.02	0.22	0.33
schooling	8.6	4.5	0.0	18.0
SD INPRES pgm exposure	0.64	0.82	0.00	6.40
wages	12.3	1.1	6.9	16.0
water supply and sanitation	0.43	0.20	0.00	2.35
woman	0.34	0.48	0.00	1.00
year of birth	1965	8	1950	1975
Sample 2 (N=4023)				
father's schooling	5.2	3.9	0	18
proportion enrolled in 1971	0.17	0.06	0.03	0.61
proportion of children in 1971	0.27	0.02	0.22	0.33
schooling	8.5	4.6	0.0	18.0
SD INPRES pgm exposure	0.54	0.75	0.00	6.40
wages	12.3	1.1	6.9	16.0
water supply and sanitation	0.43	0.20	0.00	2.35
woman	0.35	0.48	0.00	1.00
year of birth	1965	7	1950	1975

Table 33. Mean programme exposure and schooling by SD INPRES programme exposure

	SD INPRES programme		
Mean	not exposed	exposed	difference
SD INPRES exposure	0	1.1	1.1*
schooling	7.4	9.5	2.1*
observations	1323	1823	-

Note: *statistically different at one percent.

Table 34. Schooling regressions

	Sample 1			*Sample 2*		
Dependent variable: schooling	(1)	(2)	(3)	(4)	(5)	(6)
SD INPRES pgm exposure	0.5954***	0.9561***	0.9631***	0.5683***	0.7624***	0.7635***
	(0.1956)	(0.2310)	(0.2305)	(0.1777)	(0.2310)	(0.2307)
SD INPRES pgm exposure*father's schooling	-0.1033***	-0.1065***	-0.1062***	-0.1080***	-0.1070***	-0.1068***
	(0.0255)	(0.0228)	(0.0228)	(0.0207)	(0.0208)	(0.0208)
father's schooling	0.5627***	0.5624***	0.5609***	0.5539***	0.5510***	0.5505***
	(0.0322)	(0.0322)	(0.0324)	(0.0295)	(0.0295)	(0.0295)
woman	-1.2734***	-1.2441***	-1.2139***	-1.1259***	-1.085***	-1.0743***
	(0.1985)	(0.1939)	(0.1941)	(0.1709)	(0.1717)	(0.1734)
woman*SD INPRES pgm exposure	0.6539***	0.6398***	0.6287***	0.5837***	0.5545***	0.5505***
	(0.2182)	(0.2158)	(0.2165)	(0.1985)	(0.1984)	(0.1985)
district	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y
fee abolition		Y	Y		Y	Y
part-time	Y	Y		Y	Y	
F test of excluded instruments (F statistic/P-value)	12.53/0.000	17.05/0.000	17.13/0.000	14.76/0.000	16.02/0.000	15.98/0.000
Observations	3146	3146	3146	4023	4023	4023
R-squared	0.244	0.254	0.254	0.235	0.246	0.246

Robust standard errors clustered at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round. Sample 1 is the main sample, sample 2 includes individuals who were older than 7 in the year SD INPRES schools were built in their district.

Table 35. Ordinary least squares and second-stage regressions (sample 1)

Dependent variable: monthly wages (natural log)	*Sample 1*					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
schooling	0.1049***	0.1042***	0.1029***	0.1331***	0.1351***	0.1377***
	(0.0055)	(0.0056)	(0.0054)	(0.0486)	(0.0476)	(0.0513)
father's schooling	0.0215***	0.0222***	0.0255***	0.0076	0.0070	0.0084
	(0.0054)	(0.0053)	(0.0052)	(0.0257)	(0.0249)	(0.0266)
woman	-0.5626***	-0.5783***	-0.6416***	-0.5270***	-0.5397***	-0.5993***
	(0.0475)	(0.0456)	(0.0465)	(0.0683)	(0.0682)	(0.0691)
woman*SD INPRES pgm exposure	0.0411	0.0542	0.0753*	0.0238	0.0350	0.0541
	(0.0384)	(0.0371)	(0.0388)	(0.0459)	(0.0466)	(0.0488)
district	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y
fee abolition		Y	Y		Y	Y
part-time	Y	Y		Y	Y	
Hansen J statistic (overidentification test of all instruments, P-value)				0.179	0.895	0.791
Observations	3146	3146	3146	3146	3146	3146
R-squared	0.579	0.585	0.564	0.487	0.493	0.465

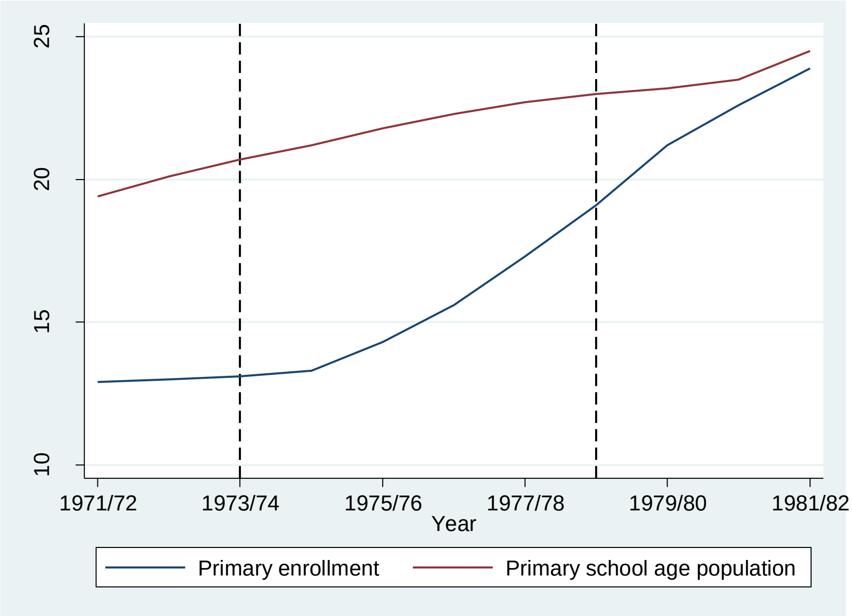
Robust standard errors clustered at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round. Excluded instruments are SD INPRES programme exposure and an interaction of SD INPRES programme and father's schooling. Sample 1 is the main sample, sample 2 includes individuals who were older than 7 in the year SD INPRES schools were built in their district.

Table 36. Ordinary least squares and second-stage regressions (sample 2)

	Sample 2					
	OLS			IV		
Dependent variable: monthly wages (natural log)	(1)	(2)	(3)	(4)	(5)	(6)
schooling	0.1037***	0.1031***	0.1027***	0.1367***	0.1317***	0.1373***
	(0.0043)	(0.0044)	(0.0044)	(0.0444)	(0.0465)	(0.0503)
father's schooling	0.0231***	0.0240***	0.0266***	0.0069	0.0099	0.0097
	(0.0044)	(0.0044)	(0.0043)	(0.0234)	(0.0241)	(0.0257)
woman	-0.5282***	-0.5407***	-0.5990***	-0.4921***	-0.5103***	-0.5626***
	(0.0464)	(0.0451)	(0.0465)	(0.0599)	(0.0612)	(0.0622)
woman*SD INPRES pgm exposure	0.0218	0.0335	0.0536	0.0046	0.0192	0.0364
	(0.0352)	(0.0340)	(0.0351)	(0.0400)	(0.0409)	(0.0423)
district	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y
fee abolition		Y	Y		Y	Y
part-time	Y	Y		Y	Y	
Hansen J statistic (overidentification test of all instruments, P-value)				0.221	0.922	0.913
Observations	4023	4023	4023	4023	4023	4023
R-squared	0.573	0.579	0.560	0.477	0.487	0.460

Robust standard errors clustered at district level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round. Excluded instruments are SD INPRES programme exposure and an interaction of SD INPRES programme and father's schooling. Sample 1 is the main sample, sample 2 includes individuals who were older than 7 in the year SD INPRES schools were built in their district.

Figure 14. Primary enrollment and school-age population, 1971-1982

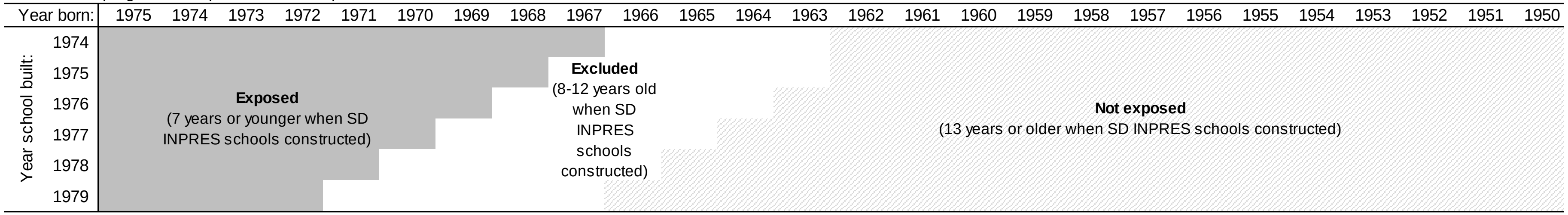


Note: School-age population is children aged 7-12 years.
Source: (World Bank 1975, World Bank 1989).

Figure 15. Distribution of schooling by gender, 2000

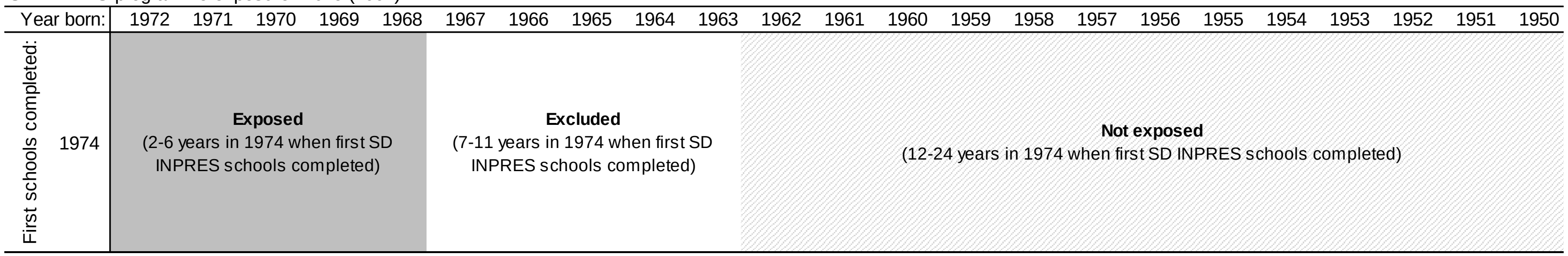
Figure 16. SD INPRES programme exposure by study

SD INPRES programme exposure: This chapter



Note: Programme intensity defined as cumulative number of SD INPRES schools constructed per 1000 children aged 5-14 in district of birth in the year in which an individual starts school.

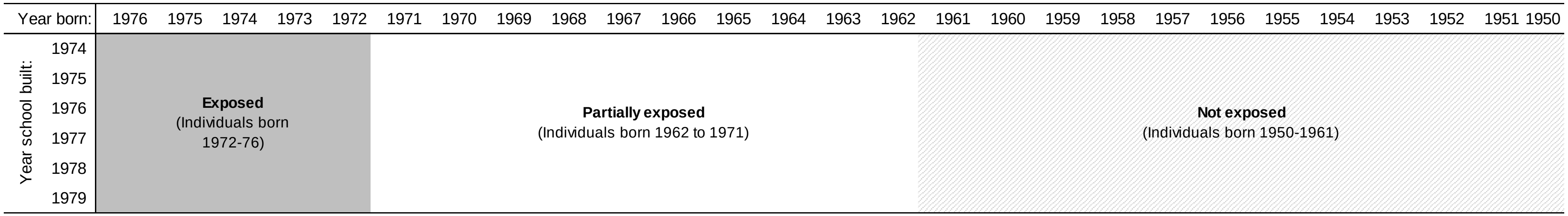
SD INPRES programme exposure: Duflo (2001)



Note: Total number of SD INPRES schools constructed between 1973/74-1978/79 in district of birth per 1000 children aged 5-14 years in 1971.

Source: Based on information provided in Duflo (2001).

SD INPRES programme exposure: Hertz and Jayasundera (2007)



Note: Average number of SD INPRES schools constructed per 1000 children aged 5-14 in district of birth in the year in which an individual starts school.

Source: Based on information provided in Hertz and Jayasundera (2007).

Bibliography

AKEE, R., SIMEONOVA, E., COPELAND, W., ANGOLD, A. and COSTELLO, J., 2010. *Does More Money Make You Fat? The Effects of Quasi-Experimental Income Transfers on Adolescent and Young Adult Obesity.* IZA Discussion Paper No. 5135. Bonn: Institute for the Study of Labor.

ANGELES, G., DIETRICH, J., GUILKEY, D., MANCINI, D., MROZ, T., TSUI, A. and YU ZHANG, F., 2001. *A Meta-Analysis of the Impact of Family Planning Programs on Fertility Preferences, Contraceptive Method Choice and Fertility.* MEASURE Evaluation Working Papers WP-01-30. Chapel Hill, NC: Carolina Population Center University of North Carolina at Chapel Hill.

ANGELES, G., GUILKEY, D., K. and MROZ, T., A., 2005. The Effects of Education and Family Planning Programs on Fertility in Indonesia. *Economic Development and Cultural Change,* **54**(1), pp. 165-201.

ANGELES, G., GUILKEY, D., K. and MROZ, T., A., 1998. Purposive Program Placement and the Estimation of Family Planning Program Effects in Tanzania. *Journal of the American Statistical Association,* **93**(443), pp. 884-899.

ARENDT, J.N., 2005. Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification. *Economics of Education Review,* **24**(2), pp. 149-160.

ASHENFELTER, O., HARMON, C. and OOSTERBEEK, H., 1999. A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias. *Labour Economics,* **6**(4), pp. 453-470.

AULD, M.C. and SIDHU, N., 2005. Schooling, Cognitive Ability and Health. *Health Economics,* **14**(10), pp. 1019-1034.

AVERETT, S., L., SIKORA, A. and ARGYS, L., M., 2008. For Better or Worse: Relationship Status and Body Mass Index. *Economics and Human Biology,* **6**(3), pp. 330-349.

AZIZ, I., J., 1990. INPRES' Role in the Reduction of Interregional Disparity. *Asian Economic Journal,* **4**(2), pp. 1-27.

BAPPENAS, 1973, 1974, 1975, 1976, 1978. *Pedoman Pelaksanaan Bantuan Pembagunan Sekolah Dasar. Menurut Instruksi Presiden Republik Indonesia.* Jakarta: Indonesian Ministry of Development Planning.

BARNUM, H., 1988. *Interaction of Infant Mortality and Fertility and the Effectiveness of Health and Family Planning Programs.* World Bank Policy, Planning and Research Working Paper 65. Washington, DC: World Bank.

BARRO, R., J., 2001. Human Capital and Growth. *American Economic Review,* **91**(2), pp. 12-17.

BARRO, R. and LEE, J., 2010. *A New Data Set of Educational Attainment in the World, 1950-2010.* NBER Working paper No. 15902. Cambridge, MA: National Bureau of Economic Research.

BECKER, G., S., 1965. A Theory of the Allocation of Time. *The Economic Journal,* **75**(299), pp. 493-517.

BECKER, G., S., 1964. *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education.* First edn. Cambridge, MA: National Bureau of Economic Research.

BECKER, G., S., 1960. An Economic Analysis of Fertility. In: A. COALE J., ed, *Demographic and Economic Change in Developed Countries.* Princeton: Princeton University Press, pp. 209-231.

BECKER, G., S. and LEWIS, G., H, 1973. On the Interaction Between Quantity and Quality of Children. *Journal of Political Economy,* **81**(2), pp. 279-288.

BEDI, A., S. and GASTON, N., 1999. Using Variation in Schooling Availability to Estimate Educational Returns for Honduras. *Economics of Education Review,* **18**(1), pp. 107-116.

BEEBY, C.E., 1979. *Assessment of Indonesian education: A guide in planning* Wellington: New Zealand Council for Educational Research.

BEHRMAN, J., R., 1997. Conceptual and Measurement Issues. In: J. BEHRMAN R. and N. STACEY, eds, *The Social Benefits of Education.* Ann Arbor: University of Michigan Press, pp. 17-67.

BEHRMAN, J., R. and DEOLALIKAR, A.B., 1995. Are There Differential Returns to Schooling by Gender? The Case of Indonesian Labour Markets. *Oxford Bulletin of Economics and Statistics,* **57**(1), pp. 97-117.

BEHRMAN, J., R. and DEOLALIKAR, A.B., 1991. School Repetition, Dropout, and the Rates of Return to Schooling: The Case of Indonesia. *Oxford Bulletin of Economics and Statistics,* **54**(4), pp. 467-480.

BEHRMAN, J., R. and STACEY, N., 1997. *The Social Benefits of Education.* First edn. Ann Arbor: University of Michigan Press.

BENEFO, K. and SCHULTZ, T.P., 1996. Fertility and Child Mortality in Côte d'Ivoire and Ghana. *The World Bank Economic Review,* **10**(1), pp. 123-158.

BEN-PORATH, Y., 1980. Child Mortality and Fertility: Issues in the Demographic Transition of a Migrant Population. In: R.A. EASTERLIN, ed, *Population and Economic Change in Developing Countries.* Chicago: University of Chicago Press, pp. 151-208.

BEN-PORATH, Y., 1976. Fertility Response to Child Mortality: Micro Data from Israel. *Journal of Political Economy,* **84**(4), pp. S163-S178.

BEN-PORATH, Y., 1967. The Production of Human Capital and the Life Cycle of Earnings. *The Journal of Political Economy,* **75**(4), pp. 352-365.

BERGER, M., C. and LEIGH, J., PAUL, 1989. Schooling, Self-Selection, and Health. *The Journal of Human Resources,* **24**(3), pp. 433-455.

BIRDSALL, N., 1998. Economic Approaches to Population Growth. In: H. CHENERY and T.N. SRINIVASAN, eds, *Handbook of Development Economics.* First edn. Amsterdam: Elsevier Science Publishers B.V., pp. 477-542.

BIRDSALL, N., 1980. *Population and Poverty in the Developing World.* World Bank Staff Working Paper SWP-404. Washington, DC: World Bank.

BOUND, J., JAEGER, D., A. and BAKER, R.M., 1995. Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak. *Journal of the American Statistical Association,* **90**(430), pp. 443-450.

BPS 2009a, 2009-last update, Infant Mortality Rate (IMR) and Under Five Mortality Rate by Province 1971, 1980, 1990, 1994, 1997, 1998 and 1999 [Homepage of Badan Pusat Statistik Republik Indonesia (Statistics Indonesia)], [Online]. Available: http://dds.bps.go.id/eng/tab_sub/view.php?tabel=1&daftar=1&id_subyek=12&notab=5 [06/05, 2011].

BPS 2009b, 2009-last update, Total Fertility Rate (TFR) by Province 1971, 1980, 1985, 1990, 1991, 1994, 1998, and 1999 [Homepage of Badan Pusat Statistik Republik Indonesia (Statistics Indonesia)], [Online]. Available: http://dds.bps.go.id/eng/tab_sub/view.php?tabel=1&daftar=1&id_subyek=12&notab=7 [06/05, 2011].

BPS, 1992. Other Proximate Determinants of Fertility. *Indonesia Demographic and Health Survey 1991.* Jakarta: Indonesia Central Bureau of Statistics, National Family Planning Coordinating Board, Indonesia Ministry of Health and Macro International Inc.

BREIEROVA, L. and DUFLO, E., 2003. *The Impact of Education on Fertility and Child Mortality: Do Fathers Really Matter Less than Mothers?* OECD Development Centre Working Paper No. 217. Paris: Organisation for Economic Co-operation and Development.

CAIRNS SINQUEFIELD, J. and SUNGKONO, B., 1979. Fertility and Family Planning Trends in Java and Bali. *International Family Planning Perspectives,* **5**(2), pp. 43-58.

CARD, D., 1999. The Causal Effect of Education on Earnings. In: O. ASHENFELTER and D. CARD E., eds, *Handbook of Labor Economics, Volume 3.* Amsterdam: Elsevier Science B.V., pp. 1801-1863.

CARD, D., 2001. Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica,* **69**(5), pp. 1127-1160.

CARNEIRO, P., LOKSHIN, M., RIDAO-CANO, C. and UMAPATHI, N., 2011. *Average and Marginal Returns to Upper Secondary Schooling in Indonesia.* Worl Bank Policy Research Working Paper 5878. Washington, DC: Word Bank.

CHERNICHOVSKY, D. and MEESOOK, O.A., 1985. *School Enrollment in Indonesia.* World Bank Staff Working Papers Number 746. Washington, DC: World Bank.

CHERNICHOVSKY, D. and MEESOOK, O.A., 1981. *Regional Aspects of Family Planning and Fertility Behavior in Indonesia.* World Bank Research Working Paper 462. Washington, DC: World Bank.

CHOPRA, M., GALBRAITH, S. and DARNTON-HILL, I., 2004. A Global Response to a Global Problem: the Epidemic of Overnutrition. *Bulletin of the World Health Organization,* **80**(12), pp. 952-958.

CLARK, D. and ROYER, H., 2010. *The Effect of Education on Adult Health and Mortality: Evidence from Britain.* NBER Working Paper No. 16013. Cambridge, MA: National Bureau of Economic Research.

CLELAND, J., 2002. *Education and Future Fertility Trends, With Special Reference to Mid-Transitional Countries.* Special Issue Nos. 48/49. New York: Department of Economic and Social Affairs, Population Division, United Nations.

CLELAND, J., BERNSTEIN, S., EZEH, A., FAUNDES, A., GLASIER, A. and INNIS, J., 2006. Family planning: the unfinished agenda. *The Lancet,* **368**(9549), pp. 1810-1827.

CONTI, G. and HECKMAN, J., 2010. Understanding the Early Origins of the Education-Health Gradient : A Framework That Can Also Be Applied to Analyze Gene -Environment Interactions. *Perspectives on Psychological Science,* **5**(5), pp. 585-605.

COURTEMANCHE, C., 2009. Longer Hours and Larger Waistlines? The Relationship Between Work Hours and Obesity. *Forum for Health Economics & Policy,* Article 5.

CUTLER, D., M. and LLERAS-MUNEY, A., 2006. *Education and Health: Evaluating Theories and Evidence.* NBER Working Paper No. 12352. Cambridge, MA: National Bureau of Economic Research.

DAROESMAN, R., 1972. Finance of Education. Part II. *Bulletin of Indonesian Economic Studies,* **8**(1), pp. 32-68.

DAS GUPTA, M., BOONGARTS, J. and CLELAND, J., 2011. *Population, Poverty, and Sustainable Development. A Review of the Evidence.* World Bank Policy Research Working Paper No. 5719. Washington, DC: World Bank.

DAVIDSON, R. and MACKINNON, J., G., 1993. *Estimation and Inference in Econometrics.* New York: Oxford University Press.

DEARDEN, L., 1999. The Effects of Families and Ability on Men's Education and Earnings in Britain. *Labour Economics,* **6**(4), pp. 551-567.

DEVEREUX, P.J. and FAN, W., 2011. Earnings Returns to the British Education Expansion. *Economics of Education Review,* **30**(6), pp. 1153-1166.

DHANANI, S. and ISLAM, I., 2004. *Indonesian Wage Structure and Trends, 1976-2000.* Report. Jakarta: International Labour Organization.

DOR, A., FERGUSON, C., LANGWITH, C. and TAN, E., 2010. *A Heavy Burden: The Individual Costs of Being Overweight and Obese in the United States.* Department of Health Policy Research Report, September. Washington, DC: The George Washington University School of Public Health and Health Services.

DUFLO, E., 2001. Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review,* **91**(4), pp. 795-813.

EIDE, E., R. and SHOWALTER, M., H., 2011. Estimating the Relation Between Health and Education: What Do We Know and What Do We Need to Know? *Economics of Education Review,* **30**(5), pp. 778-791.

FIELDS, G., S., 2008. *A Brief Review of the Literature on Earnings Mobility in Developing Countries.* ILR School Working Papers No. 101. Ithaca, NY: Cornell University.

FRANKENBERG, E., KAROLY, L., A., GERTLER, P., ACHMAD, S., AGUNG, I.G.N., HATMADJI, S.H. and SUDHARTO, P., 1995. *The 1993 Indonesian Family Life Survey: Overview and Field Report.* RAND Working Paper No. DRU-1195/1. Santa Monica, CA: RAND.

FREEMAN, R., 1986. Demand for Education. In: O. ASHENFELTER and R. LAYARD, eds, *Handbook of Labor Economic, Volume 1.* First edn. Amsterdam: Elsevier Science Publishers BV, pp. 357-386.

GERTLER, P., J. and MOLYNEAUX, J., W., 1994. How Economic Development and Family Planning Programs Combined to Reduce Indonesian Fertility. *Demography,* **31**(1), pp. 33-63.

GRABNER, M., J., 2009. *The Causal Effect of Education on Obesity: Evidence from Compulsory Schooling Laws.* PhD. Baltimore, MD: University of Maryland.

GRAHAM, C. and FELTON, A., 2005. *Variance in Obesity Across Cohorts and Countries: A Norms-Based Explanation Using Happiness Surveys.* CSED Working Paper No. 42. Washington, DC: The Brookings Institution.

GRILICHES, Z., 1977. Estimating the Returns to Schooling: Some Econometric Problems. *Econometrica,* **45**(1), pp. 1-22.

GROGGER, J., 1990. A Simple Test for Exogeneity in Probit, Logit, and Poisson Regression Models. *Economics Letters,* **33**(4), pp. 329-332.

GROSSMAN, M., 2006. Education and Nonmarket Outcomes. In: E. HANUSHEK ALAN and F. WELCH, eds, *Handbook of the Economics of Education, Volume 1.* Amsterdam: Elsevier B.V., pp. 577-633.

GROSSMAN, M., 1972. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy,* **80**(2), pp. 223-255.

GROSSMAN, M. and KAESTNER, R., 1997. Effects of Education on Health. In: J. BEHRMAN R. and N. STACEY, eds, *The Social Benefits of Education.* Ann Arbor: University of Michigan Press, pp. 69-123.

GUNTHER, I. and FINK, G., 2011. *Water and Sanitation to Reduce Child Mortality. The Impact and Cost of Water and Sanitation Infrastructure.* World Bank Policy Research Working Paper 5618. Washington, DC: World Bank.

HAYES, A., LEWIS, G. and VOGEL, R., 2003. *The National Family Planning Program in Indonesia: Review of Past Achievements, Future Directions.* Mimeo. Jakarta: Sustaining Technical Achievements in Reproductive Health Program, National Family Planning Coordinating Board.

HEROLD, P., KHAN, Q. and REIFF, H., 1983. *Public Expenditure on Education in Countries of Asia and the Pacific. Goals, Trends (1970-1980) and Future Prospects.* ED-8VWS/4. Paris: United Nations Educational, Scientific and Cultural Organization.

HERTZ, T. and JAYASUNDERA, T., 2007. *School Construction and Intergenerational Mobility in Indonesia.* American University, Department of Economics Working Papers No. 2007-18. Washington, DC: American University.

IMBENS, G., W. and ANGRIST, J., D., 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica,* **62**(2), pp. 467-476.

JOSHI, S., 2011. *How Effective Are Family-Planning Programs at Improving the Lives of Women? Some Perspectives From a Vast Literature.* Unpublished draft.

JÜRGES, H., REINHOLD, S. and SALM, M., 2011. Does Schooling Affect Health Behavior? Evidence from the Educational Expansion in Western Germany. *Economics of Education Review,* **30**(5), pp. 862-872.

KENKEL, D. and LILLARD, D., 2006. The Roles of High School Completion and GED Receipt in Smoking and Obesity. *Journal of Labor Economics,* **24**(3), pp. 635-660.

LA FERRARA, E., CHONG, A. and DURYEA, S., 2008. *Soap Operas and Fertility: Evidence from Brazil.* Bureau for Research and Economic Analysis of Development Working Paper No. 172. Durham, USA: Duke University.

LEE, B.S. and SCHULTZ, T.P., 1982. Implications of Child Mortality Reductions for Fertility and Population Growth in Korea. *Journal of Economic Development,* **7**(1), pp. 21-44.

LLOYD, C.B. and IVANOV, S., 1988. The Effects of Improved Child Survival on Family Planning Practice and Fertility. *Studies in Family Planning,* **19**(3), pp. 141-161.

LONG, J.S. and FREESE, J., 2006. *Regression Models for Categorical Dependent Variables Using Stata.* Second. College Station, Texas: Stata Press.

MALUCCIO, J., A., 1998. *Endogeneity of Schooling in the Wage Function.* FCND Discussion Papers No. 54. Washington, DC: International Food Policy Research Institute.

MASON, A., 2001. *Population Change and Economic Development in East Asia. Challenges Met, Opportunities Seized. (Contemporary Issues in Asia and Pacific).* Stanford, California: Stanford University Press.

MAULDIN, W.P., BERELSON, B. and SYKES, Z., 1978. Conditions of fertility decline in developing countries, 1965-75. *Studies in Family Planning,* **9**(5), pp. 89-147.

MAUSKOPF, J. and WALLACE, T.D., 1984. Fertility and Replacement: Some Alternative Stochastic Models and Results for Brazil. *Demography,* **21**(4), pp. 519-536.

MCCRARY, J. and ROYER, H., 2011. The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth. *The American Economic Review,* **101**(8), pp. 158-195.

MCKENZIE, D., 2005. Measuring Inequality with Asset Indicators. *Journal of Population Economics,* **18**(2), pp. 229-260.

MCNICOLL, G. and SINGARIMBUN, M., 1983. *Fertility Decline in Indonesia: Analysis and Interpretation.* Washington, D.C.: National Academy Press.

MILLIGAN, K., MORETTI, E. and OREOPOULOS, P., 2004. Does Education Improve Citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics,* (9-10), pp. 1695.

MINCER, J., 1958. Investment in Human Capital and Personal Income Distribution. *The Journal of Political Economy,* **66**(4), pp. 281-302.

MOLYNEAUX, J., W. and GERTLER, P., J., 2000. The Impact of Targeted Family Planning Programs in Indonesia. *Population and Development Review, Supplement: Population and Economic Change in East Asia,* **26**, pp. 61-85.

MONTEIRO, C., A., MOURA, E., C., CONDE, W.L. and POPKIN, B.M., 2004. Socioeconomic Status and Obesity in Adult Populations of Developing Countries: A Review. *Bulletin of the World Health Organization,* **82**(12), pp. 940-946.

OEY-GARDINER, M., 1991. Gender Differences in Schooling in Indonesia. *Bulletin of Indonesian Economic Studies,* **27**(1), pp. 57-79.

OLSEN, R., J., 1980. Estimating the Effects of Child Mortality on the Number of Births. *Demography,* **17**(4), pp. 429-444.

ORAZEM, P., F. and KING, E., M., 2008. Schooling in Developing Countries: The Roles of Supply, Demand and Government Policy. In: T.P. SCHULTZ and J. STRAUSS, eds, *Handbook of Development Economics, Volume 4.* First edn. Amsterdam: Elsevier B.V., pp. 3475-3559.

PHILIPSON, T. and POSNER, R., 2008. *Is the Obesity Epidemic a Public Health Problem? A Decade of Research on the Economics of Obesity.* NBER Working Paper No. 14010. Cambridge, MA: National Bureau of Economic Research.

PITT, M., M., ROSENZWEIG, M., R. and GIBBONS, D., M., 1993. The Determinants and Consequences of the Placement of Government Programs in Indonesia. *The World Bank Economic Review,* **7**(3), pp. 319-348.

POPKIN, B., M., 1999. Urbanization, Lifestyle Changes and the Nutrition Transition. *World Development*, **27**(11), pp. 1905-1916.

PÖRTNER, C.C., BEEGLE, K. and CHRISTIAENSEN, L., 2011. *Family Planning and Fertility. Estimating Program Effects Using Cross-Sectional Data.* World Bank Policy Reseach Working Paper No. 5812. Washington, DC: World Bank.

PRITCHETT, L., H., 1994. Desired fertility and the impact of population policies. *Population and Development Review*, **20**(1), pp. 1-55.

REINHOLD, S. and JÜRGES, H., 2010. Secondary School Fees and the Causal Effect of Schooling on Health Behavior. *Health Economics,* **19**(8), pp. 994-1001.

ROGERS, E.M., VAUGHAN, P.W., SWALEHE, R.M.A., RAO, N., SVENKERUD, P. and SOOD, S., 1999. Effects of an Entertainment-education Radio Soap Opera on Family Planning Behavior in Tanzania. *Studies in Family Planning,* **30**(3), pp. 193-211.

ROSENZWEIG, M.R. and SCHULTZ, T.P., 1989. Schooling, Information and Nonmarket Productivity: Contraceptive Use and Its Effectiveness. *International Economic Review*, **30**(2), pp. 457-477.

ROSENZWEIG, M., R. and K. I WOLPIN, KENNETH, I., 1982. Governmental Interventions and Household Behavior in a Developing Country. *Journal of Development Economics,* **10**(2), pp. 209-225.

ROSENZWEIG, M., R. and SCHULTZ, T.P., 1985. The supply of and demand for births: Fertility and its life cycle consequences. *The American Economic Review,* **75**(5), pp. 992-1015.

ROSENZWEIG, M., R. and SCHULTZ, T.P., 1983. Consumer Demand and Household Production: The Relationship Between Fertility and Child Mortality. *The American Economic Review,* **73**(2), pp. 38-42.

ROSENZWEIG, M., R. and WOLPIN, K., I., 1986. Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions. *The American Economic Review,* **76**(3), pp. 470-482.

ROSENZWEIG, M., R. and WOLPIN, K., I., 1982. Governmental Interventions and Household Behavior in a Developing Country. *Journal of Development Economics,* **10**(2), pp. 209-225.

ROSENZWEIG, M., R. and WOLPIN, K., I., 1980. Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica,* **48**(1), pp. 227-240.

SANCHEZ, H., 1979. *Community water supply and sanitation, Indonesia, September 1974 - April 1979.* SEA/EH/228. Geneva: World Health Organization.

SANDER, W., 1992. The Effect of Women's Schooling on Fertility. *Economics Letters,* **40**(2), pp. 229-233.

SCHULTZ, T.P., 2008. Population Policies, Fertility, Women's Human Capital and Child Quality. In: T.P. SCHULTZ and J. STRAUSS, eds, *Handbook of Development Economics.* Amsterdam: Elsevier B.V., pp. 3249-3303.

SCHULTZ, T.P., 2005. *Fertility and Income.* Yale University Economic Growth Center Discussion Paper No. 925. New Haven, CT: Yale University.

SCHULTZ, T.P., 1997. Demand for Children in Low Income Countries. In: M. ROSENZWEIG R. and O. STARK, eds, *Handbook of Population and Family Economics.* 1A edn. Elsvier Science Publishers B.V., pp. 349-430.

SCHULTZ, T.P., 1994. Human Capital, Family Planning, and Their Effects on Population Growth. *American Economic Review,* **84**(2), pp. 255-260.

SCHULTZ, T.P., 1988. Education Investments and Returns. In: H. CHENERY and T.N. SRINIVASAN, eds, *Handbook of Development Economics, Volume 1.* First edn. Amsterdam: Elsevier Science Publishers B. K, pp. 543-630.

SCHULTZ, T.P., 1978. Fertility and Child Mortality over the Life Cycle: Aggregate and Individual Evidence. *American Economic Review,* **68**(2), pp. 208-215.

SCHULTZ, T.P. and JOSHI, S., 2007. *Family Planning as an Investment in Development: Evaluation of a Program's Consequences in Matlab, Bangladesh.* Yale University Economic Growth Center Discussion Paper No. 951. New Haven, CT: Yale University.

SOBAL, J., RAUSCHENBACH, B. and FRONGILLO, E., A., 2003. Marital Status Changes and Body Weight Changes: a US Longitudinal Analysis. *Social Science & Medicine,* **56**(7), pp. 1543-1555.

SOBAL, J. and STUNKARD, A., J., 1989. Socioeconomic Status and Obesity: A Review of the Literature. *Psychological Bulletin,* **105**(2), pp. 260-275.

STAIGER, D. and STOCK, J.H., 1997. Instrumental Variables Regression with Weak Instruments. *Econometrica,* **65**(3), pp. 557-586.

STRAUSS, J., BEEGLE, K., SIKOKI, B., DWIYANTO, A., HERAWATI, Y. and AND WITOELAR, F., 2004. *The Third Wave of the Indonesia Family Life Survey: Overview and Field Report, Volume I.* RAND Labor and Population Working Paper Series WR144/1-NIA/NICHD. Santa Monica, CA: RAND.

STRAUSS, J. and THOMAS, D., 1995. Human Resources: Empirical Modeling of Household and Family Decisions. In: J. BEHRMAN R. and T.N. SRINIVASAN, eds, *Handbook of Development Economics, Volume 3A.* First edn. Amsterdam: Elsevier Science B.V., pp. 1883-2023.

THOMAS, D., 1999. Fertility, Education and Resources in South Africa. In: C. BLEDSOE H., J. CASTERLINE B., J. JOHNSON-KUHN A. and J. HAAGA G., eds, *Critical Perspectives on Schooling and Fertility in the Developing World.* Washington, DC: National Academy, pp. 138-180.

TWD, 2009. *Indonesia Contextual Analysis in Water Supply and Sanitation Sector.* Draft Report. The Water Dialogues Indonesia.

UNESCO, 2008. *Education for All by 2015. Will We Make It?* EFA Global Monitoring Report 2008. Geneva: United Nations Educational, Scientific and Cultural Organization.

UNESCO, 2004. *Gender and Education for All: The Leap to Equality.* EFA Global Monitoring Report 2003/04. Geneva: United Nations Educational, Scientific and Cultural Organization.

UNESCO, 1976. The Indonesian 'INPRES' Primary School Building Programme. *Bulletin of the UNESCO Regional Office for Education in Asia,* **4**(17), pp. 147-164.

UNITED NATIONS, 2010-last update, World Marriage Patterns 2000 [Homepage of Department of Economic and Social Affairs, Population Division, United Nations], [Online]. Available: http://www.un.org/esa/population/publications/worldmarriage/worldmarriage.htm [12/20, 2011].

UNITED NATIONS, 2002. *Population Bulletin of the United Nations: Completing the Fertility Transition.* Special Issue Nos. 48/49. New York: Department of Economic and Social Affairs, Population Division, United Nations.

UNITED NATIONS ENVIRONMENT PROGRAM, 1989. *Environmental Data Report.* GEMS Monitoring and Assessment Research Centre, Oxford: Basil Blackwell.

WARWICK, D., P., 1986. The Indonesian Family Planning: Government Influence and Client Choice. *Population and Development Review,* **12**(3), pp. 453-490.

WEBBINK, D., MARTIN, N., G. and VISSCHER, P., M., 2010. Does Education Reduce the Probability of Being Overweight? *Journal of Health Economics,* **29**(1), pp. 29-38.

WHO 2011a, 12/01, 2011-last update, Global Database on Body Mass Index [Homepage of World Health Organization], [Online]. Available: http://apps.who.int/bmi/index.jsp [12/01, 2011].

WHO 2011b, 2011-last update, Global Health Observatory Data Repository [Homepage of World Health Organization], [Online]. Available: http://apps.who.int/ghodata/ [12/01, 2011].

WHO, 2009. *Global Health Risks. Mortality and Burden of Disease Attributable to Selected Major Risks.* Geneva: World Health Organization.

WHO, 2004. *Global Strategy on Diet, Physical Activity and Health.* Geneva: World Health Organization.

WHO, 1998. *Obesity: Preventing and Managing the Global Epidemic.* Report on a WHO Consultation, Technical Report Series No. 894. Geneva: World Health Organization.

WHO EXPERT CONSULTATION, 2004. Appropriate Body-Mass Index for Asian Populations and Its Implications for Policy and Intervention Strategies. *The Lancet,* **363**(9403), pp. 157-163.

WITHROW, D. and ALTER, D., A., 2011. The Economic Burden of Obesity Worldwide: A Systematic Review of the Direct Costs of Obesity. *Obesity Reviews,* **12**(2), pp. 131-141.

WITLOEAR, F., STRAUSS, J. and SIKOKO, B., 2011. *Socioeconomic Success and Health in Later Life. Evidence from the Indonesia Family Life Survey.* RAND Working Paper Series No. WR-704. Santa Monica, CA: RAND.

WOLPIN, K., I., 1997. Determinants and Consequences of the Mortality and Health of Infants and Children. In: M.R. ROSENZWEIG and O. STARK, eds, *Handbook of Population and Family Economics, Volume 1A.* Amsterdam: Elsevier Science Publishers B.V., pp. 483-557.

WORLD BANK 2011a, 04/14, 2011-last update, Health, Nutrition and Population Statistics [Homepage of World Bank], [Online]. Available: http://go.worldbank.org/N2N84RDV00 [12/01, 2011].

WORLD BANK 2011b, 2011-last update, Open Data [Homepage of World Bank], [Online]. Available: http://data.worldbank.org [12/20, 2011].

WORLD BANK, 1989. *Indonesia. Basic Education Study.* 7841-IND. Washington, DC: World Bank.

WORLD BANK, 1984. *Indonesia. Policies and Prospects for Economic Growth and Transformation.* 5066-IND. Washington, DC: World Bank.

WORLD BANK, 1975. *Indonesia. Education Sector Survey Report.* 443a-IND. Washington, DC: World Bank.

WORLD RESOURCES INSTITUTE, 1988. *World Resources 1988–89.* New York: Basic Books.

Annex

Annex table 1. SD INPRES programme exposure and wage employment regressions by gender

Dependent variable: wage employment	*Women*			*Men*		
(excluded category: self-employed)	(1)	(2)	(3)	(4)	(5)	(6)
SD INPRES pgm exposure	-0.1346	-0.0961	-0.1061	0.0803	0.2385**	0.2451**
	(0.099)	(0.121)	(0.128)	(0.058)	(0.107)	(0.114)
SD INPRES pgm exposure*father's schooling	-0.0024	-0.0037	-0.0014	-0.0142**	-0.0192**	-0.0207**
	(0.010)	(0.009)	(0.009)	(0.006)	(0.008)	(0.009)
father's schooling	0.0533***	0.0496***	0.0503***	0.0561***	0.0556***	0.0612***
	(0.011)	(0.015)	(0.015)	(0.008)	(0.007)	(0.007)
province	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y
part-time	Y	Y		Y	Y	
Observations	2,766	2,766	2,766	3,349	3,349	3,349
Pseudo R-squared	0.110	0.131	0.117	0.165	0.177	0.155

Probit. Robust standard errors clustered at province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round.

Annex table 2. SD INPRES programme exposure and occupation regressions by gender

Dependent variable: occupation category	*Women*						*Men*					
(excluded category: blue collar)	services	white collar	services	white collar	services	white collar	services	white collar	services	white collar	services	white collar
Relative risk ratios	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SD INPRES pgm exposure	-0.1630	-0.3178	-0.6185	-0.3742	-0.5226	-0.4134	0.2886	0.4949**	0.3188	0.4496*	0.3218	0.4493*
	(0.301)	(0.317)	(0.436)	(0.553)	(0.423)	(0.551)	(0.199)	(0.226)	(0.310)	(0.240)	(0.311)	(0.231)
SD INPRES pgm exposure*father's schooling	-0.0172	-0.0046	-0.0638***	-0.0045	-0.0508**	-0.0032	-0.0347*	-0.0296*	-0.0507**	-0.0423**	-0.0514**	-0.0405**
	(0.039)	(0.020)	(0.019)	(0.033)	(0.020)	(0.033)	(0.019)	(0.017)	(0.024)	(0.017)	(0.024)	(0.017)
father's schooling	0.1412***	0.2422***	0.1398***	0.2452***	0.1424***	0.2425***	0.1233***	0.1413***	0.1317***	0.1460***	0.1328***	0.1414***
	(0.054)	(0.028)	(0.052)	(0.033)	(0.051)	(0.032)	(0.022)	(0.024)	(0.025)	(0.023)	(0.025)	(0.023)
province	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971			Y	Y	Y	Y			Y	Y	Y	Y
year of birth*water supply and sanitation			Y	Y	Y	Y			Y	Y	Y	Y
part-time	Y	Y	Y	Y			Y	Y	Y	Y		
Observations	927	927	927	927	927	927	1,918	1,918	1,918	1,918	1,918	1,918
Pseudo R-squared	0.210		0.266		0.251		0.093		0.121		0.118	

Multinomial logistic regressions. Robust standard errors clustered at province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round.

Annex table 3. SD INPRES programme exposure and primary school completion by gender

Dependent variable: primary school completion (excluded category: less than primary school completion)	*Women*			*Men*		
	(1)	(2)	(3)	(4)	(5)	(6)
SD INPRES pgm exposure	-0.1534	-0.1035	-0.1061	0.1316	0.0514	0.0592
	(0.145)	(0.289)	(0.283)	(0.117)	(0.107)	(0.106)
SD INPRES pgm exposure*father's schooling	-0.0074	-0.0219	-0.0222	-0.0106	-0.0044	-0.0046
	(0.018)	(0.022)	(0.023)	(0.018)	(0.021)	(0.021)
father's schooling	0.2607***	0.2757***	0.2766***	0.1542***	0.1486***	0.1501***
	(0.017)	(0.016)	(0.017)	(0.023)	(0.021)	(0.021)
province	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y
part-time	Y	Y		Y	Y	
Observations	1,076	1,076	1,076	2,051	2,051	2,051
Pseudo R-squared	0.458	0.503	0.502	0.242	0.272	0.270

Probit. Robust standard errors clustered at province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round.

Annex table 4. SD INPRES programme exposure and secondary school completion by gender

Dependent variable: secondary school completion (excluded category: less than secondary school completion)	*Women* (1)	(2)	(3)	*Men* (4)	(5)	(6)
SD INPRES pgm exposure	0.0816	0.1405	0.1406	0.1873	0.3729**	0.3734**
	(0.122)	(0.121)	(0.122)	(0.145)	(0.151)	(0.151)
SD INPRES pgm exposure*father's schooling	-0.0179	-0.0416*	-0.0417*	-0.0081	-0.0134	-0.0135
	(0.018)	(0.021)	(0.021)	(0.020)	(0.022)	(0.022)
father's schooling	0.2719***	0.3011***	0.3012***	0.1710***	0.1683***	0.1685***
	(0.041)	(0.041)	(0.041)	(0.026)	(0.027)	(0.027)
province	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y
part-time	Y	Y		Y	Y	
Observations	1,079	1,079	1,079	2,056	2,056	2,056
Pseudo R-squared	0.417	0.417	0.417	0.208	0.231	0.231

Probit. Robust standard errors clustered at province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round.

Annex table 5. SD INPRES programme exposure and wage employment regressions by socioeconomic status

Dependent variable: wage employment	*low SES*			*middle SES*			*high SES*		
(excluded category: self-employed)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SD INPRES pgm exposure	-0.1618	-0.1493	-0.1482	-0.0679	-0.0882	-0.0923	0.0282	0.3315***	0.3372***
	(0.170)	(0.295)	(0.293)	(0.066)	(0.072)	(0.069)	(0.034)	(0.120)	(0.097)
woman	-0.3398**	-0.3692***	-0.4166***	-0.4134***	-0.4097***	-0.4942***	-0.3689***	-0.3642***	-0.4882***
	(0.135)	(0.143)	(0.151)	(0.086)	(0.088)	(0.090)	(0.082)	(0.084)	(0.075)
woman*SD INPRES pgm exposure	-0.0135	-0.0028	-0.0168	-0.1498**	-0.1472**	-0.1389**	0.0050	-0.0286	0.0153
	(0.103)	(0.101)	(0.100)	(0.072)	(0.071)	(0.070)	(0.076)	(0.091)	(0.095)
province	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y		Y	Y
part-time	Y	Y		Y	Y		Y	Y	
Observations		1,568			3,487			1,055	
Psedudo R-squared	0.113	0.137	0.121	0.132	0.139	0.112	0.182	0.238	0.210

Probit. Robust standard errors clustered at province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round.

low SES = father no schooling, middle SES = father some or complete primary schooling, high SES = father secondary school or more

Annex table 6. SD INPRES programme exposure and occupation regressions by socioeconomic status

Dependent variable: occupation category	*low SES*		*middle SES*						*high SES*		
(excluded category: blue collar)	services	white collar	services	white collar	services	white collar	services	white collar	white collar	white collar	white collar
Relative risk ratios	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
SD INPRES pgm exposure	0.1627	0.6045**	-0.1221	0.2059	0.1511	0.1467	0.1760	0.1613	0.0029	-0.2100	-0.1854
	(0.796)	(0.271)	(0.242)	(0.275)	(0.279)	(0.304)	(0.297)	(0.289)	(0.199)	(0.340)	(0.344)
woman	-0.2306	0.1426	0.3063	0.3728**	0.2903	0.3401**	0.1772	0.4094**	0.7762*	1.0260**	1.1873***
	(0.289)	(0.204)	(0.217)	(0.182)	(0.244)	(0.168)	(0.245)	(0.189)	(0.408)	(0.427)	(0.402)
woman*SD INPRES pgm exposure	-0.3243	-0.7606**	-0.2046	-0.1995	-0.1539	-0.2367	-0.1110	-0.2584	0.0305	-0.1121	-0.1760
	(0.515)	(0.318)	(0.198)	(0.146)	(0.257)	(0.160)	(0.250)	(0.167)	(0.234)	(0.284)	(0.290)
province	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971					Y	Y	Y	Y		Y	Y
year of birth*water supply and sanitation					Y	Y	Y	Y		Y	Y
part-time	Y	Y	Y	Y	Y	Y			Y	Y	
Observations	584		1,626						503		
Psedudo R-squared	0.232		0.096		0.128		0.120		0.126	0.224	0.216

Multinomial logistic regressions. Robust standard errors clustered at province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round.

low SES = father no schooling, middle SES = father some or complete primary schooling, high SES = father secondary school or more

Annex table 7. SD INPRES programme exposure and primary school completion by socioeconomic background

Dependent variable: primary school completion		*low SES*			*middle SES*	
(excluded category: less than primary school completion)	(1)	(2)	(3)	(4)	(5)	(6)
SD INPRES pgm exposure	-0.0394	0.1729	0.1736	-0.1591	-0.1749	-0.1568
	(0.290)	(0.379)	(0.384)	(0.112)	(0.111)	(0.108)
woman	-0.9718***	-1.0811***	-1.0895***	-0.5147***	-0.5172***	-0.5644***
	(0.083)	(0.105)	(0.108)	(0.141)	(0.140)	(0.141)
woman*SD INPRES pgm exposure	0.2441**	0.3358**	0.3281**	0.3193	0.3124	0.3417
	(0.112)	(0.140)	(0.143)	(0.205)	(0.206)	(0.209)
province	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y
part-time	Y	Y		Y	Y	
Observations		630			1,794	
Psedudo R-squared	0.279	0.351	0.350	0.160	0.181	0.153

Probit. Robust standard errors clustered at province level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round.

low SES = father no schooling, middle SES = father some or complete primary schooling.

Annex table 8. SD INPRES programme exposure and secondary school completion by socioeconomic background

Dependent variable: secondary school completion	*low SES*			*middle SES*		
(excluded category: less than secondary school completion)	(1)	(2)	(3)	(4)	(5)	(6)
SD INPRES pgm exposure	-0.1311	0.1152	0.1258	0.0630	0.1575*	0.1556*
	(0.223)	(0.335)	(0.337)	(0.093)	(0.081)	(0.083)
woman	-0.9270***	-1.4006***	-1.3487***	-0.2814**	-0.2755**	-0.2972**
	(0.274)	(0.393)	(0.375)	(0.132)	(0.124)	(0.129)
woman*SD INPRES pgm exposure	-0.0552	0.1514	0.1379	0.1140	0.1077	0.1197
	(0.203)	(0.279)	(0.287)	(0.163)	(0.166)	(0.170)
province	Y	Y	Y	Y	Y	Y
year of birth	Y	Y	Y	Y	Y	Y
year of birth*proportion children 1971	Y	Y	Y	Y	Y	Y
year of birth*proportion enrolled 1971		Y	Y		Y	Y
year of birth*water supply and sanitation		Y	Y		Y	Y
part-time	Y	Y		Y	Y	
Observations		630			1,796	
Psedudo R-squared	0.304	0.438	0.433	0.078	0.110	0.109

Probit. Robust standard errors clustered at provinc level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant term and controls for survey round.

low SES = father no schooling, middle SES = father some or complete primary schooling

Annex figure 1. Map of Indonesia

Source: (Frankenberg, Karoly et al. 1995).