Essays on Civil War, HIV/AIDS, and Human capital in Sub-Saharan African Countries

Eric Djimeu Wouabe

Université d'Auvergne, Clermont-Ferrand 1
Faculté des Sciences Economiques et de Gestion
Ecole Doctorale des Sciences Economiques, Juridiques et de Gestion
Centre d'Etudes et de Recherches sur le Développement International (CERDI)

Essays on Civil War, HIV/AIDS, and Human capital in Sub-Saharan African Countries

Thèse Nouveau Régime

Présentée et soutenue publiquement le 12 Janvier 2011
Pour l'obtention du titre de Docteur ès Sciences Economiques

Par

Eric DJIMEU WOUABE

Sous la direction de

M. le Professeur Jean-Louis ARCAND

Membres du Jury :

M. Jean-Louis COMBES	Président	Professeur, Université d'Auvergne Clermont-Fd 1
M. Christian MORRISSON	Rapporteur	Professeur Emérite, Université Paris I Panthéon Sorbonne
Mme Mbolatiana RAMBONILAZA	Rapporteur	Charge de Recherches HDR, CEMAGREF, Bordeaux
M. Dean KARLAN	Rapporteur	Professeur, Université de Yale
M. Damien de WALQUE	Rapporteur	Senior Economist, DECRG, The World Bank
M. Gilles CARBONNIER	Suffragant	Professeur, Institut de Hautes Etudes Internationales et du Développement, Genève
M. João SOUSA ANDRADE	Suffragant	Professeur, Université de Coimbra
M. Jean-Louis ARCAND	Directeur	Professeur, Institut de Hautes Etudes Internationales et du Développement, Genève

A mes parents
Emmanuel Wouabé et Monique Tchimtchoua

REMERCIEMENTS

Je tiens à exprimer ma profonde gratitude aux nombreuses personnes qui de près ou de loin, et de diverses manières, ont permis la réalisation de cette thèse.
Je remercie particulièrement mon directeur de thèse, le Professeur Jean-Louis Arcand pour l'encadrement, l'attention et le soutien dont j'ai bénéficié tout au long de ces années de thèse. Sa disponibilité, ses conseils et orientations, en me laissant une certaine autonomie ont été déterminants pour la rédaction de cette thèse. Son implication actif et constant pour financer cette dernière m'ont permis d'avoir de bonnes conditions de travail pendant ces années qui y ont été consacrées.

Je remercie les Professeurs Christian Morrisson et Dean Karlan, les Docteurs Mbolatiana Rambonilaza et Damien de Walque qui ont accepté de rapporter cette thèse. Je souhaite également remercier les Professeurs Gilles Carbonnier et João Sousa Andrade d'en être les suffragants. Enfin, je remercie le Professeur Jean-Louis Combes d'avoir accepté présider le jury de cette thèse.

Mes remerciements vont aussi à toute l'équipe d'enseignants chercheurs du CERDI, aux participants des séminaires de l'école doctorale, ainsi qu'aux participants des conférences au cours desquelles j'ai présenté mes travaux ; pour l'ensemble de leurs remarques et suggestions. Je pense particulièrement aux Professeurs Elisabeth Sadoulet, Alain de Janvry, Dan Levy, Robert Lalonde et à deux referees anonymes du Journal Health Economics. Je remercie Jean-François maystadt et Dimico Arcangelo. Je remercie également l'ensemble du personnel administratif du CERDI pour leurs appuis logistiques durant ces années de thèse. Je pense notamment à Jaco, Martine, Annie, Michelle et Odette.

Le dernier chapitre de cette thèse a nécessité une enquête de terrain au Cameroun. Pour avoir rendu cette enquête possible et nous donner la possibilité d'évaluer les enseignements en matière de VIH/SIDA, Je suis profondément reconnaissant au Comité National de Lutte contre le VIH/SIDA du Cameroun. En particulier, je remercie le Docteur Maurice Fezeu pour son accueil et son soutien indéfectible durant tout ce processus. Je remercie également les responsables du ministère des enseignements secondaires et en particulier Désiré Aroga et Elisabeth Tenlep. Je remercie l'ensemble des proviseurs, directeurs des établissements présents dans notre échantillon dont la coopération nous a permis de réaliser cette enquête. Je n'oublierai pas de remercier les élèves et enseignants qui ont bien voulu répondre à nos questions.

Je remercie également le Professeur Adama Diaw pour son appui. Enfin, je remercie le Global Developement Network pour avoir financé cette enquête.

J'adresse par ailleurs mes remerciements à mes collègues et amis de Clermont-Ferrand qui m'ont permis de passer d'agréables moments pendant ces années de thèse. Je pense particulièrement à Luc Désiré Omgba, Calvin Djiofack, Yougbaré Lassana, Guirane Ndiaye, Maliki Amadou, Chrystelle Tsafack, Fousseini Traoré, Léandre Bassolé, Marie-Charlotte Buisson, Fred Aubery, Samba Mbaye, Christian kouontchou, Amadou Bachir Diallo, Jackie Makalle, Carole Yimga, Lancine Condé, Thierry Kangoye, Christian Ebeké, Florent Kuitché, Florentin Djiengoué, Clarisse Nguedam Ntouko, Lassina Traoré.

Je tiens à remercier les membres de l'Association Camerounaise des Etudiants et Anciens Etudiants de Clermont-Ferrand (ACEAC) pour leur amitié. Je pense notamment à Eric Gabin, Chloé, Rodrigue, Herman, Hugues, Isidore, Sébastien, Franck, Guy.

Je remercie mes amis, qui même n'étant pas à Clermont-Ferrand, m'ont témoigné tout au long de ces années de thèse leurs soutiens. Je pense particulièrement à Alain Kounga, Eric Lucas Fotso, Christophe Belle Nseke, Justin Nounamo Fotué, Aimé Raherindraibe, Emilie Feutcha et Sandrine Tchana.

Je remercie Gabriel Tourek pour la relecture de l'introduction et Dominique Ongolo pour son assistance quant à la mise en forme.

Tous mes remerciements au Docteur Albert Zeufack pour des nombreux conseils qu'il m'a prodigués durant ses années de thèse. Je tiens également à remercier Liliane Nguela Guedem pour son soutien, ses encouragements lors des moments de doute.

Enfin, Je tiens à exprimer toute ma reconnaissance et mes remerciements à mes parents, Emmanuel Wouabé et Monique Tchimtchoua pour leur amour, leur patience, leur soutien indéfectible durant toutes ces longues années d'études. Je leur dédie cette thèse. Je remercie mes frères et soeurs qui m'ont toujours accordé sans aucune réserve leur soutien moral, très important dans les moments de doute et de découragement. En particulier, je remercie ma grande sœur Solange Tchouelag Wouabé pour son attention, son soutien permanent, ses encouragements constants tout au long de ma scolarité. Je tiens aussi à saluer la mémoire de mes grands parents, Joseph Djimo et Lucas Diefé et mon oncle Youta Moise. Je témoigne enfin ma profonde gratitude à mes grandes mères Lucienne Minché et Justine Kembou.

ABSTRACT

This thesis is based on three essays. The first chapter analyses the impact of 27 years of civil war in Angola on human capital, expenditures per adult equivalent and fertility. The prediction of the effects of civil war is done through a neoclassical unitary household model in the tradition of Rosenzweig. Using instrumental variable method, this thesis shows that civil war has a negative and disastrous impact in short-term on health of children, this effect is persistent. Civil war has no impact on expenditures per adult equivalent. It increases enrollment and decreases fertility in the short term. The second chapter of this thesis analyzes the effectiveness of a social program in a conflict country such as Angola and explores whether this effectiveness depends on the intensity of the conflict. Our identification strategy is based on the political geography of the deployment of the program based on a model of spatial competition of Hotelling. This thesis shows that the Angola Social Fund had a positive impact on expenditures per adult equivalent and on one of the main anthropometric measurements namely the height for age z-score. The program's effectiveness in function to the intensity of the conflict is analyzed using the local instrumental variable estimator. The thesis shows that the program's effectiveness increases with the intensity of the conflict. The last chapter of this thesis analyzes in the case of Cameroon, the impact of teacher training on HIV/AIDS. The two criteria for selecting participating schools, leads us to choose as identification strategy the regression discontinuity design. This thesis shows that 15 to 17 year old girls in teacher training schools are between 7 and 10 percentage points less likely to have started childbearing. For 12 to 13 year old girls, the likelihood of self-reported abstinence and condom use is also significantly higher in treated schools.

Keywords: Angola, Civil war, Instrumental variable, Local instrumental variable, Anthropometric, Cameroon, HIV/AIDS, Regression discontinuity design.

RÉSUMÉ

Cette thèse s'articule autour de trois essais. Le premier chapitre analyse l'impact des 27 ans de guerre civile en Angola sur les dépenses par équivalent adulte, le capital humain et la fécondité. La prédiction des effets de la guerre se fait à l'aide d'un modèle néoclassique de ménage unitaire dans la tradition de Rosenzweig. A partir de l'approche d'estimation par variable instrumentale, cette thèse montre que la guerre civile a un impact négatif et désastreux à court terme sur la santé des enfants, cet effet est persistant. La guerre civile n'a pas d'impact sur les dépenses par équivalent adulte. Elle accroit la scolarisation et décroit la fécondité à court terme. Le second chapitre de cette thèse analyse l'efficacité d'un programme social dans un environnement en conflit comme celui de l'Angola et s'interroge si cette efficacité dépend de l'intensité du conflit. Notre stratégie d'identification est basée sur la géographie politique du déploiement du programme basée sur un modèle de compétition spatiale à la Hotelling. Cette thèse montre que le Fond Social Angolais a eu un impact positif sur les dépenses par équivalent adulte et sur l'une des principales mesures anthropométriques à savoir le z-score de la taille pour âge. L'efficacité du programme en fonction de l'intensité du conflit est analysée à l'aide de l'estimateur de variable instrumental local. La thèse montre que l'efficacité du programme augmente avec l'intensité du conflit. Le dernier chapitre de cette thèse analyse dans le cas du Cameroun, l'impact de la formation des enseignants en matière de VIH/SIDA. Les deux critères retenus pour le choix des écoles participant au programme, nous amène à choisir comme stratégie d'identification la régression discontinue. Cette thèse montre que les filles âgées de 15 à 17 ans dans les écoles traitées sont moins susceptibles d'avoir une grossesse involontaire. Pour les élèves âgés de 12 à 13 ans, la probabilité d'abstinence auto déclarée et l'utilisation du préservatif est également significativement plus élevé dans les écoles traitées.

Mots-clefs : Angola, Guerre civile, Estimation par variable instrumentale, Estimateur de variable instrumental local, Anthropométrie, Cameroun, VIH/SIDA, Modèle de Régression discontinue.

Contents

List of Tables

List of Figures

General Introduction

Among the challenges faced by sub-Saharan Africa countries in its development process, two stand as major impediments: civil war and HIV/AIDS.

Over the last 50 years nearly 20 African countries –about 40 percent of sub-Saharan Africa countries –have experienced at least one period of civil war. Nearly a third of countries in sub-Saharan African had active civil wars or conflicts during the mid-1990s (Blattman and Miguel 2008). Out of 30 major conflicts recorded in 2000, ten of the twenty-three civil wars were in Sub-Saharan Africa. In 2000, these wars are estimated to have resulted in over four million deaths and to have cost affected countries more than $138 billion in U.S. dollars (Bodea and Elbadawi 2008).

Civil war is defined as an internal armed conflict causing more than 1,000 battle deaths in a single year(Gleditsch, Wallensteen, Eriksson, Sollenberg, and Strand 2002). Civil war is accompanied by the loss of life directly related to war, degradation of living conditions, malnutrition and disease exacerbated by war, destruction of physical goods, large movement of people, suffering and trauma of any kind with long term effects, family disintegration, loss of household assets, and the destruction of social capital. Altogether, this translates into a lower level of production in agriculture, industry, and declines in trade. At the state level, civil war leads to a budget deficit due to a combination of increased military spending and lower state revenue. As a consequence, expenditures in education and health decline and chronic inflation ensues, most particularly when the deficit is financed by seigniorage. Here we see a story behind the declines in economic growth, low economic development and high poverty rates that characterize most African countries in south of Sahara.

In view of these trends, Rodrik (1999) argues that domestic social conflicts are a key to understanding why growth rates lack persistence and why so many countries have experienced a growth collapse since the mid-1970s. Collier (1999) shows that during civil war the annual growth rate is reduced by 2.2 percent in the developing countries. A fifteen year cycle of civil war would thus reduce per capita GDP by around 30 per cent. More specifically, Bodea and Elbadawi (2008) show that sub-Saharan Africa has been disproportionately impacted by civil war, which explains a substantial share of its economic decline, including a widening income gap relative to East Asia. Civil wars have also been costly for sub-Saharan Africa. In the case of Sudan, a typical large African country experiencing a long-duration conflict, the cost of war amounts to $46 billion in U.S. dollars, which is roughly double the country's current stock of external debt.

Similarly, HIV/AIDS poses an unprecedented development and human challenge, especially in sub-Sahara Africa. In many countries, the epidemic has cut life expectancy and robbed society of millions of people in their prime working years. It has dimmed the hope of living full and productive lives for unimaginable numbers of infants, children, and young adults.

Currently, more than 36 million people worldwide are living with HIV/AIDS, including 2.3 million children under the age of 15 (UNAIDS 2008). An estimated 24.2 million (67% of the total global seropositive population) people are living with the disease in sub-Saharan Africa. Globally, the annual number of new HIV infections in 2007 was 2.7 million, with approximately 1.9 million becoming infected in sub-Saharan Africa. Sub-Saharan Africa accounts for 70% of the annual number of new HIV infections (UNAIDS 2008). Indeed, the disease has become the fourth leading cause of adult mortality in the world and the second in sub-Saharan Africa (WHO, 2002). According to the Joint United Nations Programme on HIV/AIDS, HIV infections are spreading most quickly within young populations. Young people aged 15–24 account for 45% of all new HIV infections in adults, and many young people still lack accurate and complete information regarding how to avoid exposure to the virus.

HIV/AIDS leads to a direct depletion in stocks of human capital, as skilled workers die prematurely and younger workers are prevented from reaching their labor productivity potentials. In addition, the disease contributes to the persistence of poverty by affecting the accumulation of human capital and has adverse effects on the nutritional status of children particularly for orphans. For instance, Case and Carly (2006) examine the impact of parental death on children's outcomes. They find the loss of a child's mother is a strong predictor of poor schooling outcomes and children whose fathers have died live in significantly poorer households, measured on a number of dimensions. The potential negative long-run impact of HIV/AIDS on economic development can be quite substantial. Bell, Bruhns, and Gersbach (2006) estimate that in Kenya by 2040, GDP per adult will be 11 percent less than it would have been in the absence of a large-scale HIV epidemic.

Theoretical studies surveyed in Haacker (2004) typically predict 1.0 percent to 1.5 percent declines in GDP growth rates for the worst affected countries (prevalence rates above 20 percent). Results on the empirical link between the epidemic and economic growth seem to be mixed(Bloom and Mahal (1997);Corrigan, Glomm, and Mendez (2005);Papageorgiou and Stoytcheva (2005);Ahuja, Wendell, and Werker (2006);Lorentzen, McMillan, and Wacziarg (2008)). Because the HIV/AIDS epidemic dramatically affects mortality rates, some authors posit that parents will choose to have more children as an "insurance policy" guaranteeing a certain number of survivors. Analysis of evidence for 44 countries in Africa concludes that HIV/AIDS affects fertility rates positively and school enrollment rates negatively, mitigating the negative effect of the epidemic on population growth and reducing the amount of human capital investment (Kalemli-Ozcan 2006). In aggregate, these mechanisms result in slower per capita economic growth.

Civil war and HIV/AIDS harm human capital, a fundamental determinant of economic growth and economic development, in the short term and in the long term. Recent literature shows that countries affected by civil wars and/or HIV/AIDS experiment lower economic performances than others, risking falls into poverty traps. Several questions associated to civil conflicts and HIV/AIDS therefore emerge.

First, why study the impact of the civil war on the human capital? In the context of post-conflict reconstruction strategies, the international community, NGOs, and state governments could increase resource allocation to sectors such as health and education, provide assets to households for the promotion of agriculture, build infrastructure devastated by civil war; alternatively, stakeholders may take use resources to take preventative action by addressing the determinants of civil wars in sub-Saharan Africa. In theory, some combination should guide policy but increasing funding for social programs or agriculture to mitigate the negative effects of civil war might not be effective if not targeted to those most affected by civil war, and post-conflict reconstruction could itself become a potential source or replicator of inequality. As Blattman and Miguel (2008) point out, the crucial question is not whether wars harm human capital stocks, but rather in what ways, how much, for whom, and how persistently. These questions are crucial for understanding war's impacts on economic growth and inequality, as well as for determining priorities for post-conflict assistance.

In this thesis, we seek to empirically validate the disastrous effects of civil war in the short-term on human capital while analyzing the long term effects of civil war on human capital. One key question is whether, for example, children born in areas of conflict even after the end of the war are affected by civil war. The answer to this question is useful for formulating effective post-conflict strategies. The Angolan case is used to address these issues. We assess the short term effects of civil war on the human capital but more importantly we study the long term effects. Another aspect of this thesis is to examine whether the introduction of programs such as a social fund are effective strategies in mitigating the effects of civil war on the human capital. This question seems relevant given the institutional vulnerabilities faced by conflict-affected and post-conflict countries and which might affect the effectiveness of the public social actions towards the poor. Countries in conflict and post-conflict countries face different challenges; to provide an inexhaustive list: these nations have weak governance, weak social capital, a lack of qualified staff, and so on. These weaknesses can limit the impact of social funds.

Despite advances in medicine, currently there is no drug to permanently cure HIV/AIDS nor a vaccine. The only way to reduce prevalence rates requires promoting preventative practices among youth, who represent the biggest proportion of new infection and who may benefit from the knowledge and reformed attitudes that lead to the reduction of risk behaviors related to HIV/AIDS. There are four main ways to reach young people through advocacy around HIV/AIDS. First, community-based programs for young people who are not in school organize awareness campaigns, commonly in the rural areas. Targeting out-of-school youth is a key component in any HIV prevention program, given that they are more likely to be sexually active than those enrolled in educational institutions. The second channel is mass media which reach youth through media outlets and social marketing campaigns. The third channel is youth service, which provides a "hook" to connect young people with HIV prevention information and services. Finally, there are school-based HIV/AIDS education programs.

Many studies attempt to assess the impact of these different awareness strategies regarding HIV/AIDS. Studies which evaluate the impact of the community-based programs show that they have determined strong behavioral change among out-of-school youth, especially when they are delivered by established youth service organizations (Maticka-Tyndale and Brouillard-Coyle 2006). According to a meta-analysis of six youth-oriented media campaigns in various low and middle income countries, this strategy had a measurable impact on HIV/AIDS-related risk behaviors (National Research Council 2005). Another meta-analysis of 22 studies of school-based HIV/AIDS prevention education programs in low- and middle-income countries shows positive effects of this program on age of first sex, frequency of sex, number of partners, use of condoms or contraceptives, and frequency of unprotected sex (WHO (2006); Biddlecom, Hessburg, Singh, Bankole, and Darabi (2007)).

However, it is very difficult to ascribe any causal change in behavior related to HIV/AIDS in all of these programs. Indeed, as noted by Duflo, Dupas, Kremer, and Sinei (2006), these assessments suffer from important methodological shortcomings, especially in terms of an identification strategy to evaluate causal impact; also, in most cases, these studies have low statistical power. Rigorous evaluation of in-school programs addressing the level of knowledge, changing attitudes and reducing risky behavior related to HIV/AIDS is important for two main reasons: First, a substantial amount of financial resources are devoted to HIV/AIDS prevention programs. Most countries (89%) indicate having integrated HIV education into their secondary school curricula, with countries in sub-Saharan Africa being most likely to do so (UNAIDS, 2008). Second, rigorous evaluations of the effectiveness of such interventions have rarely been undertaken.

Some of the most striking gaps in the empirical literature on HIV/AIDS is the lack of any studies investigating the causal determinants of the spread of the disease in sub-Saharan Africa. The main macroeconomic studies on the determinants of HIV/AIDS found a strong correlation between income inequality and gender inequality and HIV/AIDS prevalence. GNP per capita is also linked negatively to the infection, the percent of foreigners and military forces among total population are correlated positively with national HIV prevalence rate while the percentage of Muslim in total population is negatively associated with the infection rate (Over 1998). Bonnel (2000) goes beyond the previous study by taking into account the reverse causality between HIV/AIDS and these determinants using a simultaneous equation method. Using cross-sectional data for 59 developing countries in 1997, she finds that the growth rate of GDP per capita and secondary school enrolment rate were non-significant. The two variables most strongly correlated with HIV infection are indeed the percent of muslim in total population and the time since the first HIV case was reported in the country. In this model, access to information, gender inequality, ethnic fractionalization, and labor migration are also strongly correlated with HIV/AIDS prevalence. Stillwaggon (2000) focuses on the economic and biomedical environment in which HIV transmission occurs. She insists that HIV/AIDS infection is primarily an infectious disease and, as such, is fuelled by poverty, malnutrition, urbanization, and international migration. A better understanding of the difference in the spread of the HIV/AIDS epidemic between countries would require looking at the factors that usually distinguish developing countries from developed ones, rather than an attribution of this difference to behavior alone. Using multivariate ordinary least squares regression with a sample of African countries, she finds that change in calorie intake, change in the Gini coefficient, and change in urban population are positively and significantly correlated with the epidemic. Change in per capita GDP is not correlated with the epidemic. Although these previous studies show a correlation between HIV/AIDS and social economic factors, in the absence of a causal link between the spread of HIV/AIDS and these determinants, it is difficult to implement public policy based on reliably evidence that could lead reductions in HIV prevalence.

To our knowledge, there is only one microeconomic study which analyses the probability of contracting HIV/AIDS using survey data. De Walque (2006) analyzes the determinants of HIV infection and associated sexual behaviors using data from the first five Demographic and Health Surveys, which include HIV testing for a representative sample of the adult population. Four important findings emerged. First, married women who engage in extra marital sex are less likely to use condoms than single women engaging in sex. Second, having been in successive marriages is a significant risk factor, as evidenced by the rises in HIV infection and risky sexual behaviors. Contrary to prima facie evidence, education is not associated positively with HIV status though schooling is one of the most consistent predictors of behavior and knowledge: education predicts protective behaviors like condom use, use of counseling and testing, discussion among spouses, and knowledge, but it also predicts a higher level of infidelity and a lower level of abstinence. Finally, male circumcision and female genital mutilation are often associated with sexual behaviors, practices and knowledge related to AIDS. Oster (2009) is to our knowledge the only existing study which seeks to analyze the causal determinants of HIV/AIDS incidence in Africa. She studies the relationship between economic activity (specifically, exports) and the incidence of HIV in Africa and argues that there is a significant and large positive relationship between the two: a doubling of exports leads to an approximate doubling of HIV incidence. She argues that this relationship is causal and that the mechanism, in part, operates through increased movement of people along trucking routes.

Regarding civil war, it seems important to look at the commonly accepted economic factors predicting the onset of conflict. This is even more important since the other subject of this thesis is the effectiveness of social programs in a context of civil war or post-conflict. Our main question asks whether, despite the collapse of institutions and destruction of social capital necessary for their effectiveness, social programs seeking to increase access to basic services in education, health, and water supply mitigate the impact of civil war on the human capital and could be use to prevent the emergence of civil war in post-conflict countries. In attempting to resolve this question begin by moving backward to another: Why were there so many civil wars in sub-Saharan Africa since the years of independence? The analysis of the causes of civil wars took place both in a macroeconomic perspective and a microeconomic perspective. In the macroeconomic perspective, the three pioneering studies on the causes of civil war are those of Collier and Hoeffler (1998);Collier and Hoeffler (2004) and Fearon and Laitin (2003).

Collier and Hoeffler, in their seminal work on the causes of civil war in cross-country regressions – which includes social factors (such as inequality, the ethnic and religious composition of a society), history (such as the time since decolonization), and politics (such as the extent of democratic political rights) as well as economic characteristics – show that in the sub-Sahara Africa three factors matter a lot for the risk of civil war: the level of income, its rate of growth, and its structure. If a country is poor, in economic decline, and is dependent upon natural resource exports, then it faces a substantial risk that sooner or later it will experience a civil war. Typically, such a country runs a risk of around one in seven every five years. In contrast, other factors predict in the opposite direction: lower risk of civil war is associated with higher levels of secondary school attainment in the population. Meanwhile, a country's ethnic fractionalization, income inequality, and democracy are not robustly statistically significant predictors of conflict risk. Fearon and Laitin (2003) take a similar cross-country approach. They show that the current prevalence of internal war is mainly the result of a steady accumulation of protracted conflicts since the 50s and 60s rather than a sudden change associated with a new post-Cold War international system. They also find that after controlling for per capita income, more ethnically or religiously diverse countries have been no more likely to experience significant civil violence in this period. The factors that explain which countries have been at risk for civil war are not their ethnic or religious characteristics but rather the conditions that favor insurgency. These include poverty, which marks financially and bureaucratically weak states and also favors rebel recruitment, political instability, rough terrain, and large populations. The main drawback of these cross-country studies is that they only identify the robust correlation of civil war and the level of income or poverty, the rate of growth, and the natural resource endowment. To address this concern and identify the causal link between civil war and the rate of growth, Miguel, Satyanath, and Sergenti (2004) identify an exogenous source of variation in incomes in sub-Saharan Africa: rainfall shocks. Using annual rainfall growth rates as an instrumental variable for per capita income, they find that a 5 percent drop in incomes increases the likelihood of a civil conflict in the following year by nearly one half. This analysis highlights the role that income shocks play in generating armed conflict in Africa.

In a microeconomic perspective, studies on the causes of civil war aimed to answer three main questions: first, what determines the participation of individuals in an armed group?; second, what about the internal geography of countries affects the onset and the spread of civil war?; third, does the organization and conduct of civil war matter for the onset of civil war and pursuit of civil war?

Regarding the motives that determine the participation of individuals in armed conflict, the authors who use the within country regional data to explore the factors that predict violence and conflict show that there is a strong correlation between local conflict and unemployment, inequality, natural disasters, changes in sources of incomes, and clustering of ethnic groups within villages (Barron, Kaiser, and Pradhan 2004). Murshed and Gates (2005) find that the intensity of conflict across the districts of Nepal is significantly explained by the degree of inequalities. More specifically, studies investigating the characteristics of individuals directly involved in the conflict show that the probability of a household member to become a perpetrator of genocide in the case of the Rwanda is positively correlated with a poor wage worker status and landless status (Verwimp 2005). Humphreys and Weinstein (2008) in the context of Sierra Leone civil war find that poverty, a lack of access to education, and political alienation predict participation in both rebellion and counter rebellion. Factors that are traditionally seen as indicators of grievance or frustration may instead proxy for a more general susceptibility to engage in violent action or a greater vulnerability to political manipulation by elites.

Concerning the role of the internal geography of countries in the emergence and spread of civil war, the main studies are those of Buhaug and Rød (2006),which study the local determinants of African civil wars between 1970 and 2001, and find that territorial conflict is more likely in sparsely populated regions near the state border, at a distance from the capital, and without significant rough terrain. Conflict over state governance is more likely in regions that are densely populated, near diamond fields, and near the capital city.

Finally, empirical studies on the organization and lead up to civil war are very limited. According to Blattman and Miguel (2008) important areas for attention include: the choice between alternative recruitment and incentive strategies; the strategic use of civilian violence; information dissemination and repression; the dynamics of war escalation; the identity and tactics of rebel leaders; and the formation and growth of rebel organizations. The causal relation that emerges from these studies shows that the main causal factors of civil war are poverty and weak economic growth in sub-Saharan Africa countries.

In this thesis, we have used two countries, Angola and Cameroon. Angola offers a unique experience to assess the short term and long term impact of civil war on the human capital and also to see if implementing a program such as a social fund increases access to basic services which may mitigate the effects of civil war on the human capital. Whether this type of program can be used as an instrument to promote peace in a post-conflict country is also retained as a research question.

Angola, officially the Republic of Angola, is a country in south central Africa bordered by Namibia on the south, the Democratic Republic of the Congo on the north, and Zambia on the east; its west coast is on the Atlantic Ocean. Angola was a Portuguese overseas territory from the 16th century to 1975. The Angolan Civil War began in Angola after the end of the war for independence from Portugal in 1975. The war featured conflict between two primary Angolan factions, the communist *Movimento Popular de Libertação de Angola* (MPLA) and the anti-communist *União Nacional para a Independência Total de Angola (*UNITA). A third movement, *the Front de Libération de l'Etat du Cabinda* (FLEC), an association of separatist militant groups, fought for the independence of Cabinda. By the time the 27 years war was formally brought to an end in 2002. An estimated 500,000 to 1 million people had been killed, more than 4.5 million people displaced, and more than 9 million landmines used.

The Angolan Civil War was one of the largest, longest, and most prominent armed conflicts of the Cold War. Despite the fact that Angola is a country rich in natural resources, oil, and diamond in particular, at the end of the civil war in 2002 Angola was 164 out of 175 countries on the Human Development Index. Life expectancy in Angola in 2001 is 40.2 years, the adult literacy ratio is 42%, and the combined primary, secondary and tertiary gross enrolment ratio was 29%, while the respective sub-Saharan averages for these indicators were 46.5 years for life expectancy, 62.4 % for adult literacy and 44% for the combined enrolment ratio. In 2001, stunning prevalence was 45%. Population without sustainable access to an improved water source was 62%. The Angolan government decided to address the situation of extreme poverty and widespread famine conditions following an episode of peace after the Lusaka Agreement signed by the government and UNITA on November 20, 1994. With the support of international donors, mainly the World Bank, the government created in 1994 the *Fundo do Apoio Social* (FAS) with the objective of reducing extreme poverty by providing the most vulnerable populations with basic social infrastructure in health, education , and access to drinking water.

With an HIV prevalence of 5.5% among adults aged 15-49, Cameroon is, with Ivory Coast, the most affected country in West Africa (for a region that is typically less affected by the epidemic, compare to East or Southern Africa). This level of prevalence hides significant disparities between regions and between age groups. In urban areas and in the North West prevalence is up over 8%, although the figure is 2 % in northern regions (National Institute of Statistics and ORC Macro, 2004). The group hardest hit by HIV/AIDS are young women aged between 20-24 years with a prevalence rate of 7.9%. This can be explained by the fact that despite a sizeable proportion (72.1%) aware that using condoms is a means of prevention against HIV/AIDS, only 45.4% reported having used a condom at last high risk sex in 2004. The survey also showed that the proportion of young people 15-24 years that have sex with non regular partner increased in Cameroon from 2004 to 2006. This situation makes young people an ideal target for intervention because of their greater vulnerability.

Given the evolution of the pandemic, Cameroon's response was early, with the establishment of a control program against HIV/AIDS in 1986. Government action was organized around a five year national strategic plan with a component focus on the education system and was officially launched March 31, 2004 by the Minister of National Education. In this perspective, actions to reduce the prevalence of HIV/AIDS at school are: the organization of instructional days on STD and HIV/AIDS; the capacity building of stakeholders in the implementation of the PDA program (*Participation et Développement de l'Adolescence*) in six provinces, 56 institutions and 22 sites outside the school with support from UNICEF;

the caravan "No AIDS campaign"; "Campus Without AIDS"; and health clubs. In 2006, the Cameroonian government, with technical support from UNESCO and funding from the African Development Bank, decided to introduce into the education system in Cameroon a module named "EVF/EMP et VIH/SIDA". We have, in collaboration with the National AIDS Control Committee, Ministry of Health, the Ministry of Secondary Education, conducted an impact evaluation of teacher training on the knowledge, attitudes and behaviors related to HIV/AIDS among youth in Cameroon.

This thesis is based on three essays. The first chapter studies the impact of 27 years of civil war in Angola on human capital, expenditures per adult equivalent and fertility. We capture human capital through two variables – namely, anthropometric measurements including the height-for-age z-score, the weight-for-age z-score of children under 5 years and the probability of an children being enrolled in the primary school. To predict the effects of civil war on human capital, expenditures per adult equivalent and fertility, we build a unitary household model in the tradition of Rosenzweig (1990), which shows how civil war affects our outcomes. More explicitly, we assume in our model that during the civil war, we have a significant drop in wages, increased food commodity prices and a reduction in other sources of household income. Our model predicts that during the war we should have: a deterioration of the nutritional status of children, increased school attendance, and a drop in average household consumption. The impact of civil war on fertility depends upon the relative magnitudes of two effects; on the one hand, the decrease in the wage rate increases fertility; on the other hand, the decrease in household income reduces fertility, with the net effect depending upon the relative magnitudes of these two phenomena. We use the method of instrumental variables to empirically test our model. We justify our identification strategy using a theoretical model. This theoretical model falls into the rent-seeking category of explanations for civil war initiated by Grossman (1991, 1992, and 1999) and his collaborators (Gershenson and Grossman 2000). We show through this model that the localization of conflict is not exogenous but depends of the distance and richness of the area that can be correlated with our outcomes. Our model predicts that the probability of the outbreak of conflict in an area is an increasing function of wealth or payoff and that there is a decreasing and convex relationship between the probability of outbreak of conflict in a region and the distance that separates the community and the headquarters of rebel leader. The data used in this chapter comes from two main sources: *the Inquérito aos agregados familiares sobre receitas e despesas* (National Household Survey on Expenditures and Income , henceforth, IDR 2000) collected in Angola in 1999/2000, and data on civil war in Angola from the work of Ziemke (2007), who, based on archives, libraries, and news agency files (a total of 186 sources from over twenty countries), constructed a database of 9,216 individual battle and massacre events that took place in the Angolan war over a forty one year period (1961-2002), as reported in the press. In our empirical analysis we distinguish two periods that allow us to test the short term impact (1995-2000) and the long term impact (1975-1994) of civil war in Angola. The main conclusion of this chapter is that the civil war has a disastrous effect in the short term on the height-for-age z-score and the weight-for-age z-score of children less than 5 years. The long term impact of civil war shows that its effects are persistent; we find a negative impact of civil war on children who have no direct exposure to war. In the short term, the probability of a child being enrolled in primary school is greater in communities most affected by civil war. We find no long term impact of civil war on expenditure per adult equivalent and fertility. In this chapter, we show that the effect of war depreciates over time without this depreciation necessarily being linked to a reconstruction program in targeted areas of conflict. This result and, in particular, the disastrous effects of the civil war on anthropometry suggests that the establishment of emergency programs in health, nutrition, and access to drinking water during the civil war if possible and during the post-conflict recovery is important to mitigate the impact of civil war. It is significant that in this chapter, we cannot see all the effects of civil war on the human capital, expenditures per adult equivalent, or the fertility for the whole of Angola; this would require a counterfactual, a "what would have been Angola without civil war". These results imply that it is important in a post-conflict countries to design projects to increase access to basic services such as health, education, and income-generating projects that could mitigate the impact of civil war.

Having demonstrated in the first chapter the negative effects of civil war on the human capital and particularly on child anthropometry and highlighted the benefits of implementing the social programs promoting access to primary health care, the primary education, and water supply programs in post-conflict recovery, the second chapter examines three questions: whether social funds improve the health of children in Angola, if Angola social funds can also be used to mitigate the risk of relapse into civil war, and whether the effectiveness of such programs depends on the intensity of the conflict in the community participating in the program. While the allocation of projects to communities was not random, and to estimate the causal impact of Angola Social Fund on our outcomes we have constructed a political economy model allowing us to understand the deployment of a social fund during conflict. Unlike other social funds that use a poverty map to target the allocation of the project, the case of Angola shows that the purpose of program implementation was to consolidate government political support in areas contested with UNITA, the main rebel group, as well as to maintain the support of the population in areas to the rear of frontline communities. Based on a simple model of spatial competition in the tradition of Hotelling (1929), we show that the likelihood of FAS deployment should be an increasing function of the distance separating a given community from the government's forward bases in Bailundo (Huambo province), and a decreasing function of the distance to UNITA's main base at the time in Andulo (Bie province). This deployment model based our identification strategy to determine the impact of social fund on our outcomes. Data used here are those from the IDR 2000 and data on civil war in Angola come from work of Ziemke (2007). We also have data from the Angolan project management unit for communities that have received a project or not and the nature of the project received. Our results show that the Social Fund in Angola demonstrates a positive and significant impact on the height-for-age z-score for children under 5 years and though mitigates the negative effects of the civil war. We also find a positive and significant impact of social funds on expenditures per adult equivalent. The program's impact on expenditure per adult equivalent is more important with the intensity of the conflict. This result suggests that during reconstruction, programs such as social funds should primarily be focused in areas that were the most affected by conflicts. This means that at the beginning of the post-conflict reconstruction, it is important to design a map of conflict or civil war. This result shows the importance of targeting of social programs in war or in post-conflict for improving the health of children. The impact of social funds operates mainly through increasing the incomes of households living in communities participating in the program for the expenditures per adult equivalent and anthropometry. For the anthropometric measures, the impact of social funds functions also through the reduction of distance and time of access to drinking water. This result shows that despite the specific situations of countries in conflict or post-conflict, characterized by low social capital, weak rule, destruction of basic social infrastructure, social programs such as the social fund can be used as tools to mitigate the impact of civil war on human capital, especially the health of children. Social programs can also serve as instruments for generating income that can be a driver of sustainable peace in post-conflict.

The last chapter of this thesis addresses HIV/AIDS, the second challenge to human capital we identify in sub-Saharan Africa countries. In this chapter, we evaluate the effect of teacher training on the level of knowledge, attitudes, and behaviors related to HIV/AIDS among youth. The Cameroonian government sought to reduce the incidence of HIV/AIDS among young people in school and decided in 2006 to introduce into the education system a teaching module devoted to HIV/AIDS with technical support from UNESCO and financial support of the African Development Bank. During 2006-2007, about 2,000 teachers were trained to offer this new module. Teachers were trained to speak to students of various sexually transmitted diseases including HIV/AIDS, the transmission modes and manifestations of these STDs, and, more specifically, the different methods of protection against

STDs and HIV/AIDS with special attention paid to condom use. The choice of the original schools to which teachers belonged were not random but decided by the Cameroon Government in partnership with the African Development Bank (AfDB). Two criterion were used to choose the schools to participate in the program: the school must have been located in a village where the AfDB had financed the construction or renovation of a unit specializing in HIV/AIDS in the local health center; a selected village must have possessed between 1 and 4 secondary schools while schools located in villages with more than 4 schools did not participate in the program. This administrative rule that determined the deployment of the program is used as the key to our identification strategy. Our identification strategy exploits the natural experiment generated by being either just above or just below the 4 secondary schools per town threshold adopted as the key determinant of program deployment. This sharp regression discontinuity design (RDD) approach is akin to a randomization which is valid around the threshold. In the words of Lee and Lemieux (2010), we conducted a "local randomized experiment".

In this chapter, we use data that we collected in Cameroon between December 2007 and January 2008, nearly 10 months after the training of teachers. Data were collected in about 120 secondary schools in Cameroon located in 5 regions. Given the possibility of differentiating the effects of the program on students, we considered two groups of students, the youngest of average age 12-13 years and the older of average age 15-17 years. In each school we collected data at four different grades. For the sample of students aged 12-13 years, the data were collected from students attending $5^{ème}$ and $6^{ème}$; for the sample of students aged 15-17 years, data were collected from students attending Class $1^{ère}$ and 2^{nde}. We collect data on the level of knowledge, attitudes, and behaviors related to HIV/AIDS and childbearing. We also collected data on the characteristics of the school, teachers, and socio-demographic characteristics of the student's family. We find that 15 to 17 year old girls in teacher training schools are between 7 and 10 percentage points less likely to have started childbearing, an objective proxy for the incidence of unprotected sex. They are also significantly more likely to have used a condom during their last sexual intercourse. For 12 to 13 year old girls, the likelihood of self-reported abstinence and condom use is also significantly higher in treated schools, while the likelihood of having multiple partners is significantly lower. Our results show that the teaching of HIV/AIDS awareness in schools can be an effective way to limit the risk behaviors associated with the spread of the HIV/AIDS among young people in sub-Saharan Africa, particularly for countries with high school enrollment.

Chapter 1

Civil War and Human Capital in Angola

The heart's memory eliminates the bad and magnifies the good.

Gabriel Garcia Márquez, *Love in the Time of Cholera*

1.1 Introduction

Over the past forty years, Africa has suffered 398 armed conflicts, roughly one half of the world's total. Of these conflicts 125 have been civil wars, again representing roughly one half of the world's total (Gleditsch, Wallensteen, Eriksson, Sollenberg, and Strand 2002). The Angolan civil war was amongst the most brutal of these. It lasted 27 years, from independence from Portugal in 1975 to the death of UNITA leader Jonas Savimbi in 2002.[1] It involved the violent interplay of a varied and continuously shifting cast of characters, including a Marxist government born of a liberation movement (the MPLA), initally led by a poet (Agostinho Neto) and supported by the Soviet Union and its Cuban proxy, two other liberation movements (the FNLA and UNITA), the former Portuguese colonial masters, apartheid South Africa, the United States through various CIA-funded covert activities, and neighbouring countries such as Mobutu Sese Seko's Zaire.[2] According to the International Peace Research Institute of Oslo (PRIO), approximately 1.5 million people died during the conflict, while 4 million were displaced. To this day, large parts of the countryside are still peppered with landmines. A country that, on the eve of independence, boasted the largest manufacturing sector outside of apartheid South Africa and Ian Smith's Rhodesia (Zimbabwe), today has no manufacturing sector to speak of. And while Angola, in 1975, was the world's fourth largest producer of coffee; finding Angolan coffee, even in

[1] Indeed, sustained violence started as far back as 1961. For a history of the decolonization of Portugal's African colonies, see MacQueen (1997).

[2] The MPLA is the *Movimento Popular para a Liberação de Angola*; Agostinho Neto was its historic leader; he died in 1978 and was succeeded by Jose Eduardo Dos Santos, who remains president to this day, just as the party remains in power. The FNLA is the *Frente Nacional para a Liberação de Angola*. UNITA is the *União Nacional para a Independencia Total de Angola*. Cuban troops were dispatched to Angola in 1976 and remained until 1999. At times there were over 50,000 Cuban soldiers stationed in various parts of the country.

Luanda, has almost become an exercise in futility.[3]

This chapter attempts to empirically assess the impact of the Angolan conflict on household outcomes, using unique survey data collected near the end of the civil war in 2000, which we merge with extremely detailed information on the location of episodes of violence.

The research presented in this chapter is also related to a new and rapidly growing microeconomic literature on the impact of civil on human capital outcomes, particularly in the African context. The chapter contributes to this body of research in several manners. First, and contrary to most empirical work on the impact of civil war, we motivate our empirical specification with an explicit Neoclassical model of household behavior, which yields a set of testable empirical predictions.

Second, and again contrary to the bulk of the literature on civil war, we assess the *long-run* impact of the Angolan civil war on human capital. We focus on this issue because existing evidence on such long-run effects is mixed, with the notable exception of Miguel and Roland (2005). Most studies consider the long-run effects of a relatively *brief* exposure to civil war. In this chapter, we consider not only individuals who were exposed over a relatively long period, but also the impact on those who were *not* directly exposed, but who happened to live, at the time of the survey, in areas that had been severely affected in the past. These analyses allow us to highlight areas in which interventions geared towards mitigating civil war's influence could be concentrated, and to compute a rough "depreciation rate" of the long-run impact of civil war.

Third, we pair our Neoclassical household model to a rent-seeking model of civil war which shows that conflict is more likely, *ceteris paribus*, in relatively rich areas. This theoretical model of conflict also furnishes us with an identification strategy based on the distance to rebel headquarters.

Fourth, and this stems directly from the Neoclassical household refered to earlier, our analysis disaggregates human capital into its nutritional and educational components, with

[3]See Dilolwa (1978) and Dias Amaral (2004). On the colonial period, see Marques (1963).

what appear to be counter-intuitive results in terms of educational attainment, but which are readily interpretable in the context of a Neoclassical household model.

Finally, our results are based on an instrumental variables approach that allows for the intensity of civil war to be endogenous, while most existing work is based on a difference-in-differences (DID) approach.

This last point is worth stressing in that most existing studies of the effects of civil war assume that its occurrence is *exogenous*, as would be the case with random shocks due to drought. In the literature which examines the impact of rainfall on child health, most authors use a cohort approach. A typical example is Hoddinott and Kinsey (2001), who use a Zimbabwe dataset running from 1983 to 1998 and rainfall shocks in 1991-2 and 1995-6 to identify the impact of drought shocks on child height. Alderman, Hoddinott, and Kinsey (2006) take this one step further, and combine rainfall shocks with exposure to civil war (the log number of days the child was alive prior to 18 August 1980, with this variable assumed to be exogenous), alongside maternal fixed effects, to identify differences in pre-school nutritional status across syblings. They find that children exposed to such shocks are 3.4cm shorter and complete 0.85 years less of schooling.

In the case of the war in Iraq initiated in March 2003, Guerrero-Serdan (2009) analyzes a possible causal effect of the war on nutritional outcomes of children. Her results indicate that children born in areas affected by high levels of violence are 0.8 cm shorter than children born in low violence provinces. Initially proposed by Akresh and de Walque (2008) for the case of the Rwandan genocide, her approach involves constructing a pseudo-panel of children of similar age born before and after the war, and examining differences in their nutritional (or educational, for the case of Rwanda) outcomes between low and high conflict intensity provinces.

While the cohort-based DID approach allows one to control for time-invariant unobservable heterogeneity, it does conceptually treat the intensity of civil war as an exogenous occurrence, and neglects a large body of literature, such as Hagelstein (2008) for the case of Algeria, Hegre, Ostby, and Raleigh (2007) for the case of Liberia, or Buhaug and Rød (2006) for all of Africa for the 1970-2001 period, that shows that conflict intensity is positively related to either income or proxies for wealth. It is also based on the seldom-tested hypothesis of parallel trends, in which it is assumed that the evolution of child health or educational attainment would have been the same in various geographical areas, had they been subjected to identical levels of conflict intensity. Finally, it is based on the assumption that any unobservables which determine conflict intensity are time-invariant and are thus purged through the DID procedure.[4]

For the case of Angola, and given the extreme length of its civil war (27 years), no pre-1975 household datasets are, to the best of our knowledge, available. As such, the DID approach is not open to us. In contrast to the studies cited above, we therefore adopt an instrumental variables approach, geared towards controlling for unobservables that would simultaneously affect conflict intensity and household outcomes, in order to establish the magnitude of the causal relationship going from conflict intensity to human capital outcomes.

Identification of the impact of the conflict on various response variables is rendered possible because, as in many civil wars, there was a great deal of variation in conflict intensity across different regions. Our identification strategy is based on the plausibly exogenous variation in conflict intensity induced by the geographic location of a community with respect to the headquarters of UNITA, the main rebel group, as is predicted by the simple rent-seeking model of civil war that we present. Our identification strategy is in part inspired by current research in the political geography literature. Buhaug and Rød (2006), for example, examine the determinants of the spatial distribution of civil wars in Africa, using various geographically-motivated covariates. Our identification strategy is also similar in spirit to that adopted by Miguel and Roland (2005), who use the distance

[4]Note also that, in most of the papers cited above, the authors transform a continuous measure of conflict intensity into a dummy variable indicating either high or low conflict intensity. We prefer to eschew such arbitrary normalizations, and use the number of casualties within a given radius of the community in which the household lives, over a given period, as our measure of conflict intensity.

to the 17th parallel (the arbitrarily chosen border between North and South Vietnam) as a source of exogenous variation in the intensity of the American bombing campaign during the Vietnam war.

In line with the predictions of a Neoclassical unitary household model which links child health, household expenditures, school enrollment and fertility behavior, we find large short-term negative impacts of conflict intensity —corresponding to the five years (1995-2000) preceding the collection of our response variables— on child health and fertility, a large positive impact on school enrollment, and a statistically insignificant negative impact on household expenditure per adult equivalent. In contrast, the impacts are very small when we consider long term effects by using a measure of conflict intensity that corresponds to the 1975-1994 period (this period goes from the beginning of the civil war to the Lusaka Protocols, which brought a temporary end to the fighting). Entering both measures of conflict intensity simultaneously brings the difference between short- and long-run effects into even sharper focus. For child anthropometrics, both effects are negative, but with a steep dropoff in the impact of casualties which corresponds to a "depreciation rate" for casualties of roughly 63% per annum. For household expenditures and fertility, on the other hand, the short-term effects are large and negative and the long-term effects are small and positive, corresponding, respectively, to (i) a Neoclassical growth model with vintage capital and convergence effects and (ii) the historical patterns of fertility behavior during and after conflicts that have been documented by demographers. Thus, while contrary to Gabriel Garcia Márquez, time does not entirely eliminate the bad, it certainly does mitigate it (for child anthropometrics), while a small measure of good does emerge in the long-term for household expenditures and fertility.

The rest of this chapter is organized as follows. In section 1.2 we provide a brief guide to the existing literature that deals with the impact of civil war on child anthropometrics, household expenditures, school enrollment and fertility. In section 1.3,we construct a unitary household model in the tradition of Rosenzweig (1990) that provides testable hypotheses on the likely effects of conflict intensity on our four categories of response variables. In section 1.4, we present our basic empirical specification, and show why OLS based estimates of the impact of conflict intensity are likely to be biased. In section 1.5 we spell out our identification strategy and construct a theoretical model of civil war, inspired by the logistics of UNITA military operations and the rent-seeking literature, that yields a specific (quadratic) functional form that should be taken by our exclusion restrictions in the first-stage reduced forms. We then consider the likely direction of biases if our identification strategy were flawed, present our household and conflict intensity data, and report reduced form estimates for conflict intensity based on our theoretical model. Section 1.6 presents our empirical results as well as a theoretical framework within which they can be interpreted. We then present baseline instrumental variables estimates of the impact of conflict intensity on child anthropometrics, household expenditures per adult equivalent, school enrollment, and the number of live births, and highlight the conformity of the short-term results with our theoretical household model. We also explore the differences between the short- and long-run impact of conflict intensity, and document significant cross-sectional non-linearities in the relationships linking conflict intensity to our response variables using a nonparametric 2SLS procedure due to Newey and Powell (2003). Section 1.7 concludes.

1.2 Civil war and household outcomes

There is a large cross-country literature on the causes and consequences of civil war, often associated with the work of Paul Collier and his collaborators (see e.g. Collier and Hoeffler 1998, Collier 1999, Collier and Sambanis 2005). Blattman and Miguel (2008) provide a comprehensive survey of the civil war literature, and note how the cross-country approach cannot help one to understand either the mechanisms through which civil wars (and wars in general) affect households, the magnitude of these effects as a function of the severity of conflict, or their distributional consequences. Moreover, the approach is of little help in trying to design post-conflict recovery strategies. They also note that much remains to be done in terms of intra-country analyses. Buhaug and Lujala (2005) also critique the country-level approach —"ecological fallacy" is the term, due to Robinson (1950), used in the political geography literature— and plead for the use of Geographical Information Systems (GIS)-based data at the intra-country level.[5]

Recent years have seen a spectacular growth in intra-country work. Three studies of the impact of strategic bombing of countries as diverse as Japan (Davis and Weinstein 2002), Germany (Brakman, Garretsen, and Schramm 2004) and Vietnam (Miguel and Roland 2005) all show that recovery from catastrophic levels of violence is speedy. In the Japanese case, cities recovered their typical *relative* size within 15 years of the end of the conflict, despite great inter-city heterogeneity in the severity of bombing. In the German case, differences in bombing intensity had a significant but temporary impact on post-war city growth in Germany as a whole as well as in the Federal Republic separately, but not in the DDR. For the Vietnamese case, and in what is perhaps the most imaginative of these studies in terms of identification strategy, Miguel and Roland (2005) show that, 27 years after the end of the war in 1975, heterogeneity in bombing intensity did not leave significant differences in poverty rates, consumption levels, infrastructure, literacy or population density. They also provide a useful interpretation of the speed of recovery in terms of the transitional dynamics of the Neoclassical model of economic growth.

Several studies, dealing with Tajikistan, Rwanda, Burundi and Sierra Leone, have considered the impact of civil war on human capital outcomes (especially schooling and child nutritional status) by combining household survey data with GIS-based data on conflict intensity, and adopting either a difference-in-differences (DID) or instrumental variables (IV) approach. Shemyakina (2006) considers the impact of the conflict in Tajikistan during the 1992-1998 period on school enrollment using a DID approach. She finds that exposure to the conflict, as measured by past damage to household dwellings, had a large signific-

[5] Justino (2007) provides a survey of the emerging intra-country literature.

ant negative effect on the enrollment of girls, and little, or no, effect on the enrollment of boys. She also finds that girls who were of school age during the conflict and lived in conflict-affected regions were less likely to complete mandatory schooling.

Akresh and de Walque (2008) consider the impact of the 1994 Rwandan genocide on the educational outcomes of different cohorts of Rwandan children. Apart from a DID approach based on two independent surveys which straddle the genocide, they also use distance to the Ugandan border as an exclusion restriction in order to identify the impact of conflict intensity on educational outcomes. They find that children exposed to the genocide experienced a drop in educational achievement of almost one-half year of completed schooling, and were 15 percentage points less likely to complete the third or fourth grade.

For the case of Burundi, Bundervoet, Akresh, and Verwimp (2008) combine household survey data from 1998 with event data on the timing and location of armed conflicts from 1994 to 1998 to examine the impact of civil war on children's health status. The identification strategy exploits arguably exogenous variation in the war's timing across provinces and the exposure of different birth cohorts to the fighting. They find that an additional month of war exposure decreases children's height for age $z-$scores by 0.047 standard deviations compared to non-exposed children.[6]

Finally Bellows and Miguel (2006, 2008) consider the impact of the 1991-2002 Sierra Leone civil war using household survey data, and focus on household-level exposure to various acts of violence. They find that three years after the end of the civil war there are no lingering impacts of war-related violence on local socioeconomic conditions, such as per capita expenditures, the proportion of children enrolled in school, and child Body Mass Index (BMI).[7]

[6] Alderman, Hoddinott, and Kinsey (2006) examine the impact of pre-school malnutrition on subsequent human capital formation in rural Zimbabwe using a maternal fixed effects-IV estimator with a long term panel data set. Though the paper does not deal with the impact of civil war *per se*, it does use civil war (and drought) shocks to identify differences in pre-school nutritional status across siblings.

[7] They also consider the impact of heterogeneity in the exposure to violence on political mobilization and institutional change. In particular, they find that measures of local community mobilization and collective action – including the number of village meetings and the voter registration rate – are significantly higher in areas that experienced more war violence, conditional on prewar and geographic controls.

Demographers have also been active participants in the debates surrounding the impact of conflict on household behavior, though they are less sensitive to issues of identification strategy than are economists, and usually restrict their analyses to breaks in long-run trends. The standard references on 20th century developed countries in this literature, such as Rindfuss and Sweet (1977), Hobcraft (1996) or Schwartz (1997), all find a decline in fertility during war, followed by a rebound in the immediate postwar period, though it is doubtful whether these results can be generalized to developing countries that are either in the very early stages, or have not even reached, the demographic transition.

For the case of Angola, the only paper we are aware of is that by Agadjanian and Prata (2002), who use the 1996 Multiple Indicator Cluster Survey (MICS), some two years after the end of a major outbreak of fighting, and find evidence of a wartime drop and a postwar rebound in fertility, although these trends vary greatly, depending on the type and degree of exposure to war and on women's socioeconomic characteristics.

Most papers on developing countries in the demographic literature find no impact of war on an overall decline in fertility. This is true for the Lebanese civil war (Khlat, Deeb, and Courbage 1997, Kulczycki and Saxena 2000) and —for the Iran-Iraq conflict of the 1980s— for Iranian women (Ladier-Fouladi and Hourcade 1997). For Ethiopia, Lindstrom and Berhanu (1997) show that there are short-term fluctuations in fertility that are associated with political and military instability as well as famine conditions, though their data do not allow them to convincingly disentangle these effects. The same if true for the case of Eritrea (Woldemicael 2005).

1.3 A Neoclassical household model of child nutrition, school enrollment, household expenditures and fertility

Consider a simple unitary household model in the tradition of Rosenzweig (1990) in which household preferences $U(.)$ are a function of average child income i, a composite household consumption good c, time spent in school t, child nutrition x, and fertility (the number of children) n:

$$U(i, c, t, x, n).$$

In contrast to Rosenzweig (1990), and in order to allow for potentially different effects of civil war on child nutritional status and school enrollment, we allow for a separate effect on household preferences of these two variables. For simplicity, we write the returns to human capital as $\alpha h(t, x)$ where α is the corresponding rental rate. Child income is given by:.

$$i = \alpha h(t, x) + b, \tag{1.1}$$

where b represents (possibly *inter vivos*) transfers from parents to children. As in Rosenzweig (1990), i can be understood as child income in adulthood: this specification allows for the intertemporal tradeoff inherent in the investment today in a child's human capital tomorrow, while eschewing the complications that would arise if one were to explicitly formulate a two-period model of choice.

Since education t corresponds to the time that a child spends in school, gross income from the employment of children will be given by $nw(T - t)$, where w is the wage rate and T is a child's time endowment. The opportunity cost of educating children is foregone earnings nwt, whereas the unit cost of feeding one child will be denoted by p. The household's budget constraint, in which child earnings are pooled with those of the parents at the time at which schooling takes place and nutrition is supplied, is therefore given by:

$$c = y + n\left[w(T-t) - b - px\right], \tag{1.2}$$

where $w(T-t) - px$ represents the net resources brought by one child to the household, and y is exogenous parental income. The FOCs associated with this problem are given by:

$$\frac{dU}{dt} = U_i \alpha h_t - U_c n w + U_t = 0, \frac{dU}{dx} = U_i \alpha h_x - U_c n p + U_x = 0,$$
$$\frac{dU}{dn} = U_c\left[w(T-t) - b - px\right] + U_n = 0, \frac{dU}{db} = U_i - U_c n = 0.$$

Substituting the FOC in b into those in t and x then allows one to write $U_t = U_c n\left(w - \alpha h_t\right)$ and $U_x = U_c n\left(p - \alpha h_x\right)$. Combining these last two expressions with the FOC in n then yields the expressions for the marginal rates of substitution (MRS) between time in school and fertility, nutritional status and fertility, and nutrition and time in school, which must equal the corresponding ratios of shadow prices at the optimum:

$$MRS_{t,n} = \frac{U_t}{U_n} = \frac{n\left(w - \alpha h_t\right)}{b + px - w(T-t)},$$
$$MRS_{x,n} = \frac{U_x}{U_n} = \frac{n\left(p - \alpha h_x\right)}{b + px - w(T-t)},$$
$$MRS_{x,t} = \frac{U_x}{U_t} = \frac{w - \alpha h_t}{p - \alpha h_x}.$$

For illustrative purposes, let household preferences be given by a simple Cobb-Douglas specification, $U(i, c, t, x, n) = i^{\epsilon} c^{\theta} t^{\beta} x^{\gamma} n^{\delta}$, and let $\alpha h(t, x) = \alpha t$, i.e. the returns to human capital are solely a function of time spent in school. Then straightforward calculations imply that:

$$x^* = \frac{Tw\gamma}{p\left(\beta + \gamma - \delta + \epsilon\right)}, t^* = \frac{Tw\beta}{\left(w - \alpha\right)\left(\beta + \gamma - \delta + \epsilon\right)}, n^* = \frac{y\left(\beta + \gamma - \delta + \epsilon\right)}{Tw\left(\delta + \theta\right)}, c^* = \frac{y\theta}{\delta + \theta}.$$

Differentiation with respect to w, p and y then yields comparative statics results which are summarized in the first column of Table 1.4.

If the sole impact of civil war were to disrupt labor markets, thereby leading to a fall in the rental rate w of labor, our model predicts that one should observe a fall in child nutritional status, an increase in school enrollment, an increase in fertility and no change in average household consumption expenditures. In contrast, if civil war increases the price of food p, the model predicts a deterioration in the nutritional status of children but no impact on school enrollment or fertility, with consumption per capita remaining constant. Finally, if the impact of civl war were solely to decrease exogenous household income y, one would expect a fall in fertility and in average household consumption, but no change in terms of child nutritional status or school enrollment. Combining all of these changes in exogenous factors leads to three unambiguous predictions. More explicitly, if civil war decreases the wage rate, increases food prices and decreases exogenous household income, one should expect:

- a deterioration in child nutritional status;
- an increase in school enrollment;
- a decrease in average household consumption.

The impact of civil war on fertility depends upon the relative magnitudes of two effects; on the one hand, the decrease in the wage rate increases fertility; on the other hand, the decrease in household income reduces fertility, with the net effect depending upon the relative magnitudes of these two phenomena.[8]

[8] In column 2 of Table 1.4 we consider a model in which nutrition is the sole factor that goes into producing human capital, whereas column 3 reports comparative statics results for perfect substitutability between nutrition and educational attainment in the production of human capital (we also considered a Leontieff specification: results were, of course, similar to those presented in columns 1 and 2). For the latter case, the comparative statics results are, qualitatively speaking, identical to those presented in column 1. When the returns to human capital are independent of educational attainment and are solely a function of childhood nutritional status, the one notable difference in comparative statics is that civil war should have no impact on educational attainment.

1.4 Empirical specification

Let i denote children, h households and c communities; N will denote sample size. The basic structural equation that we are seeking to estimate is given by:

$$y_{ihc} = x_{ihc}\alpha + w_c\beta + \varepsilon_{ihc}, \tag{1.3}$$

where y_{ihc} is the $N \times 1$ vector associated with the outcome of interest (such as child health), x_{ihc} is a $N \times K_x$ matrix of child, household and community control variables, w_c is an $N \times 1$ measure of the intensity of conflict suffered by community c, and ε_{ihc} is a disturbance term. Our purpose is to consistently estimate β, which measures the impact of the intensity of conflict on our outcome variable. We decompose the disturbance term into two components:

$$\varepsilon_{ihc} = \lambda_c + \eta_{ihc}.$$

λ_c represents community-level unobservables that affect the outcome, while η_{ihc} are child- or household-level unobservables.

It is highly likely that OLS based estimates of (1.3) will lead to an inconsistent estimate of β because the intensity of conflict w_c is likely to be correlated with community-level unobservables λ_c. For example, the decision of government and rebel armed forces to engage in military operations in a given area are likely be correlated with community characteristics that are not adequately captured by the household- and community-level observables that are included in x_{ihc}. Estimating (1.3) with community-specific fixed effects solves this problem, but essentially throws the baby out with the bathwater in that the impact of community-level variables such as w_c can then no longer be identified in that estimation is wholly based on within-community variation.[9] Community-specific random effects are not feasible either, because the likely endogeneity of the intensity of conflict

[9] An alternative would be Hausman and Taylor (1981) type estimation, but it is very difficult to think of household-level characteristics that can be plausibly argued to be orthogonal with respect to λ_c.

variable implies that it will be correlated with the random effects. As such, the only solution is instrumental variables.

In order to get an idea of the likely direction of bias that might arise in estimating equation (1.3) by OLS, consider the following simple linear factor specification:

$$y_{ihc} = x_{ihc}\alpha + w_c\beta + \lambda_c\sigma + \eta_{ihc}, \tag{1.4}$$

$$w_c = z_c\pi + \lambda_c, \tag{1.5}$$

where z_c is a $N \times K_z$ matrix of covariates that determine the intensity of conflict, which we assume for the time being to be orthogonal with respect to λ_c and η_{ihc}, and where the scalar parameter σ determines the "degree of endogeneity" of the conflict intensity variable. If we estimate (1.4) by OLS, the estimated coefficient will be given by:

$$\beta_{OLS} = \frac{\text{cov}\left[M_x w_c, M_x y_{ihc}\right]}{\text{var}\left[M_x w_c\right]} = \beta + \sigma\frac{\text{var}\left[\lambda_c\right]}{\text{var}\left[w_c\right] + \text{var}\left[P_x w_c\right]}, \tag{1.6}$$

where $P_x = x\left(x'x\right)^{-1}x'$ is the $N \times N$ idempotent projection matrix associated with x_{ihc} and $M_x = I_N - P_x$ is its orthogonal complement (with I_N being the $N-$dimensional identity matrix).

In the context of the Angolan civil war, government and rebel forces were more likely to fight over regions of tactical or strategic importance, characteristics which are often associated with relatively high levels of development: this point will be brought home in the simple theoretical model of conflict presented below. In terms of *observable* measures of development, this is what is found, for example, by Hagelstein (2008), in his study of the Algerian civil war: levels of violence were higher in wealthier areas. The same is true in the study by Hegre, Ostby, and Raleigh (2007) of the conflict in Liberia. It is thus probably reasonable to assume that this positive correlation holds for *unobservables* as well, implying that $\sigma > 0$. As such, and considering the case in which the true impact of conflict intensity on our response variables is negative ($\beta < 0$), estimating (1.4) by OLS is likely to result in an upward bias in β_{OLS}, with the magnitude of this bias being increasing in the relative importance of the variance in community-level unobservables var$[\lambda_c]$. If this bias is sufficiently large relative to the true value of β, it may lead to β_{OLS} being *positive*. Note however that adding covariates x_{ihc} to the specification, as long as they are correlated with conflict intensity, reduces the magnitude of the bias, given that it increases var$[P_x w_c]$, where $P_x w_c$ is the value of conflict intensity predicted by a regression of w_c on x_{ihc}.

1.5 Identification strategy

1.5.1 Intuition

The identification strategy that we propose is based on the specifics of the Angolan civil war, with our excluded instrument being given by the distance separating community c from the location that was the main seat of UNITA headquarters during the conflict. The town of Jamba, in Cuando Cubango province (near the Caprivi Strip of colonial history fame), was the main UNITA headquarters from 1976 to 1991, and, to a lesser extent, an important logistical base for 24 years until its capture in late 1999. The location of UNITA headquarters and, in particular of Jonas Savimbi has often been highlighted as an important factor in the Angolan civil war.[10] Indeed, the conflict only came to an end once Jonas Savimbi was physically killed and his body displayed by government forces in the village of Lucusse (Moxico province) in July 2002.

According to the Angola Peace Monitor (2000), writing shortly after the fall of Jamba to government forces (*Forças Armadas Angolanas* —FAA) on 24 December 1999:

> Jamba was virtually created by apartheid South Africa and the American intelligence agency, the CIA, to serve as UNITA's headquarters.... In his book, *Angola's Last Best Chance for Peace*, Paul Hare describes Jamba as "a spread

[10] See Guidolin and La Ferrara (2007), for an event study of diamonds that uses the timing of his death as the key identifying variable.

out, well-organised guerrilla encampment, carefully planned and camouflaged to protect against air attacks".

In order for the distance to UNITA headquarters to constitute an admissible instrument, it must satisfy two conditions. First, conditional on the child-, household- and community-level covariates included in x_{ihc}, the distance to UNITA headquarters must be a statistically significant determinant of the intensity of conflict facing community c. There are good *a priori* reasons to expect this to be the case. In particular, the logistics of UNITA military operations, and the difficulties of moving men and materiel across Angola, implied that communities further away from their headquarters were, *ceteris paribus*, more costly targets for UNITA operations.

Second, for the distance to Jamba to constitute an admissible instrument, it must, *conditional on* x_{ihc}, be orthogonal with respect to λ_c. Several potential confounding factors must be accounted for amongst the covariates included in x_{ihc}. Indeed, the isolation of Jamba, lying in the furthest reaches of Cuando Cubango province near the border with Namibia, implies that the distance of a given community from UNITA headquarters might be inversely related to the community's remoteness, which itself might well be correlated with household- and community-level unobservables that affect the response variable(s). As such, it is important to include, amongst the covariates in x_{ihc}, variables that will control for such forms of remoteness.

1.5.2 Distance to rebel headquarters and conflict

In order to formalize the intuition behind our identification strategy, as well as to provide a theoretical basis for the specification that we should adopt for our first-stage reduced forms, we consider a simple model of civil war between a government and a rebel group. The model falls into the rent-seeking category of explanations for civil war initiated by Grossman (1991, 1992, 1999) and his collaborators (Gershenson and Grossman 2000).

The payoff to successful rebellion or government victory in a given geographical area will be denoted by X. We consider a sequential move game in which the government moves first, followed by the rebels. We solve the game by backward induction.

The power of rebel military forces will be denoted by R, whereas the power of government military forces will be denoted by G. In what follows, it is worth thinking of G as government military power above the "usual" level in a given region. *Ceteris paribus*, an increase in rebel military power R increases the likelihood of rebel victory in the case of conflict, and conversely for government military power. We summarise this by writing the probability of rebel victory as:

$$\Phi\left(R-G\right),$$

where, for illustrative purposes, $\Phi(.)$ is the *cdf* of the standard normal density.[11] The corresponding standard normal *pdf* will be denoted by $\phi(.)$. The expected payoff Y to the rebels is given by:

$$Y = X\Phi\left(R-G\right) - C(R),$$

where $C(R), C' > 0$, represents the cost to the rebels of projecting a level R of military power. In order to capture our idea that the projection of rebel military power becomes more costly as they move further away from their logistical base, we write:

$$C(R) = dR,$$

where d is the distance from their base.[12] The rebels choosing the projection of their military power in an optimal fashion in response to government actions leads to the solution

[11] The model readily generalizes to a specification in terms of a generic *cdf*.

[12] This corresponds to the loss-of-strength gradient model of conflict intensity, originally proposed by Boulding (1962) for inter-state conflict. For alternative models of civil war from the political geography literature, see, for example, Buhaug (2007).

to their optimization problem being given by:

$$R^*(X,G) = \underset{\{R\}}{\arg\max}\ \ X\Phi(R-G) - dR, \tag{1.7}$$

which yields:[13]

$$R^* = G + r(X,d),\ \text{ where }\ r(X,d) = \sqrt{2\log[X/d] - \log 2\pi}. \tag{1.8}$$

The preceding expression indicates that optimal rebel military power is increasing, one-for-one, in government military power, whereas optimal rebel power, *ceteris paribus*, is increasing in the size of the exogenously given prize X, and decreasing in the distance d from the rebel base. The basic problem facing the government is whether to deploy sufficient military force so as to deter the rebels altogether or whether to allow conflict to break out. We now proceed to show that, in equilibrium, the government's choice will be determined by the magnitude of X and the distance d of the region in question from the rebel base.

Note that the rebel group will be deterred from engaging in military operations when:

$$Y^* = X\Phi(R^* - G) - dR^* \leqslant 0,$$

where we normalize their reservation level of welfare to zero. It can be shown that the preceding inequality boils down to:

$$G \geqslant \widetilde{G} = Xd^{-1}\Phi(r(X,d)/4) - r(X,d), \tag{1.9}$$

[13]The first order condtion (FOC) associated with (1.7) is given by $X\phi(R^* - G) - d = 0$. For the FOC to be sufficient as well as necessary, it must be the case that we are on the downward-sloping portion of the *pdf* (i.e. $\phi'(R^* - G) < 0$), so as to ensure that the associated second order condition (SOC) is negative. Solving the preceding FOC yields two roots in R^*, only one of which, that given in (1.8), will satisfy the SOC. For the model to make algebriac sense (and to guarantee a non-imaginary solution), we also impose the regularity condition that $[X/d]^2 > 2\pi$.

where $\widetilde{G}$ is therefore the level of government military power that ensures deterrence.

Now consider the government's optimization program. If there is deterrence, the government receives a certain payoff $\Pi = X - \widetilde{G}$, where we normalize the government marginal (and average) cost of achieving a given level of military power to one. Substituting for the value of $\widetilde{G}$ from (1.9) then yields:

$$\Pi^D = X\left[1 - d^{-1}\Phi\left(r\left(X, d\right)/4\right)\right] + r\left(X, d\right). \qquad (1.10)$$

In the case of war, the government's optimization problem is given by:

$$\max_{\{G\}} \; X\left[1 - \Phi\left(R^* - G\right)\right] - G,$$

where R^* is given by (1.8). This yields the trivial solution $G = 0$, implying that the government's optimal payoff under war is given by:

$$\Pi^W = X\left[1 - \Phi\left(r\left(X, d\right)/4\right)\right]. \qquad (1.11)$$

Whether conflict breaks out or not then depends upon whether it is in the government's interest to deter the rebels or not, which in turn boils down to a comparison of equations (1.10) and (1.11). When $\Delta = \Pi^W - \Pi^D > 0$ there is war, otherwise, deterrence prevails.

Figure 1.1 plots the difference Δ between the government's net payoff under war with its net payoff under deterrence on the vertical axis, as a function of X and d. Two aspects of the Figure are noteworthy. First, for a given value of d, conflict is more likely as X grows. Second, for a given value of X, the likelihood of conflict is a decreasing and convex function of the distance d to rebel headquarters.

Our simple theoretical model and the illustration furnished by Figure 1.1 has at least two implications in empirical terms. First, since the model predicts that the likelihood of conflict is increasing in X, and since we are essentially interested, from an econometric standpoint, in measuring the impact of conflict on various proxies for X, our model highlights the endogeneity of conflict in the regressions that will constitute the empirical crux of this chapter.

Second, our theoretical model suggests that conflict is a *decreasing and convex* function of the distance to rebel headquarters. Letting z_c^{Jamba} denote the distance to UNITA headquarters in Jamba, our theoretical model therefore suggests that the underlying first-stage reduced form that corresponds to the structural equation specified in (1.4) should be given by:

$$w_c = x_{ihc}\gamma + z_c^{Jamba}\pi_1 + \left(z_c^{Jamba}\right)^2 \pi_2 + v_{ihc}, \tag{1.12}$$

with $\pi_1 < 0$ and $\pi_2 > 0$. Whether or not z_c^{Jamba} does provide any modicum of identification can be explicitly tested by estimating equation (1.12), and examining the statistical significance of π_1 and π_2.

1.5.3 Potential sources and direction of bias

So as to organize our thoughts, assume that our exclusion restriction is false, and that the structural equation is actually given by:

$$y_{ihc} = x_{ihc}\alpha + w_c\beta + \sigma\lambda_c + z_c\phi + \eta_{ihc}, \tag{1.13}$$

where the vector of parameters ϕ represents the degree of violation of the exclusion restrictions.[14] In the Angolan context, and if "remoteness" is associated with worse outcomes, it is reasonable to assume that each element of ϕ is positive: the greater the distance from UNITA headquarters, and therefore the lesser the degree of remoteness, the better are outcomes. If, for illustrative purposes, we were to estimate (1.13) using the distance to

[14] If we use z_c^{Jamba} and $\left(z_c^{Jamba}\right)^2$ as excluded instruments, ϕ is a 2×1 vector, with each element corresponding to the degree of violation of the exclusion restriction for each excluded instrument.

UNITA headquarters as our *sole* excluded instrumental variable, we would obtain:

$$\beta_{IV} = \frac{\text{cov}\left[M_x z_c, M_x y_{ihc}\right]}{\text{cov}\left[M_x z_c, M_x w_c\right]} = \beta + \frac{\text{var}\left[M_x z_c\right]}{\text{cov}\left[M_x z_c, M_x w_c\right]}\phi, \tag{1.14}$$

where $\frac{\text{var}[M_x z_c]}{\text{cov}[M_x z_c, M_x w_c]}$ is readily seen to be the inverse of the coefficient associated with z_c of a regression of conflict intensity w_c on z_c and x_{ihc}. If the marginal effect of distance from UNITA headquarters, conditional on x_{ihc}, is to reduce conflict intensity (which means that $\frac{\text{var}[M_x z_c]}{\text{cov}[M_x z_c, M_x w_c]} < 0$), it follows that β_{IV} will be biased downwards, leading potentially (for the case in which the true β is negative) to an overstatement of the deleterious impact of conflict intensity on outcomes. Equation (1.14) is clear: in order to minimize potential bias stemming from a violation of our exclusion restriction, (i) the distance to UNITA headquarters must be a strong predictor of conflict intensity (so that $\frac{\text{var}[M_x z_c]}{\text{cov}[M_x z_c, M_x w_c]}$ is small), whereas (ii) we must minimize the potential size of ϕ by including a rich set of covariates x_{ihc} in the structural equation. In particular, as mentioned earlier, we must control for the relative remoteness of the community: we do so by including the distance to Luanda among the covariates, as well as the distance to each community's provincial capital.[15]

1.5.4 Data

The estimates presented in this chapter are based on the 1999-2000 *Inquérito aos agregados familiares sobre despesas e receitas* ("national household survey on expenditures and incomes", henceforth, IDR 2000), which was collected in the provinces of Cabinda, Luanda, Lunda Norte, Benguela, Namibe, Huila and Cunene. Given the unstable security situation at the time, the survey is roughly representative of areas of Angola under effective government control, and this *caveat* should be kept in mind in interpreting our results.[16] The IDR

[15] If we have more than one exclusion restriction, as is the case with our quadratic specification in terms of the distance from Jamba, the corresponding expression for the bias of the IV estimate is $\beta_{IV} = \beta + \frac{\text{cov}[M_x P_z w_c, M_x z_c]}{\text{var}[M_x P_z w_c]}\phi$, where$P_z = z\left(z'z\right)^{-1} z'$.

[16] Angola is made up of 18 provinces. The survey was carried out by the *Gabinete de monitarização das condições de vida da população*, *Instituto nacional de estatística* (INE), *Ministério do planeamento* (MINPLAN).

2000 includes information on household composition, expenditures, education, health and fertility behavior. It uses a stratified sampling design in which 12 households were surveyed in a random fashion in 850 aldeias (villages) in rural areas and bairros (neighborhoods) in urban areas, in 50 comunas. Summary statistics for the IDR 2000 data are presented in Table 1.1

Our conflict intensity variable is based on the painstaking work of Ziemke (2007) who, based on archives, libraries and news agency files (a total of 186 sources from over twenty countries were involved), constructed a database of 9,216 individual battle and massacre events that took place in the Angolan war over a forty-one year period (1961-2002), as reported in the press.[17] She provides the number of victims and the geographical coordinates for each event, allowing us, for each location included in the IDR 2000 dataset (for which we obtained the coordinates using Google Earth), to calculate the number of war-related casualties within a given radius of the community, in a given year. We focus on four different radii: 1, 5, 10 and 20 kilometres. We also consider two distinct periods in the civil war, for historical reasons: (i) 1975-1994, which corresponds to a period going from the beginning of the civil war to the Lusaka peace accords (Spears 1999), which brought a temporary lull in the fighting; (ii) 1995-2000, a period which ends with the fall of Jamba to government forces. We also calculated the distances separating each community in our dataset from Jamba, Luanda and the appropriate provincial capital.

Summary statistics for the conflict intensity variables are presented in Table 1.2. On average, communities included in the sample are much further from Jamba (1,160 km) than they are from Luanda (380 km); they are also quite close to their respective provincial capitals (80 km).

[17] This work is part of the Armed Conflict Location and Events Dataset (ACLED).

1.5.5 Reduced form estimates

The results from estimating the reduced form for conflict intensity given by (1.12) are presented in Table 1.3. The left-hand portion of the table corresponds to the 1975-1994 period, while the right-hand portion of the table corresponds to the 1995-2000 period. The dependent variable is given by the total number of casualties, within various radii of the community.[18] These reduced forms correspond to the instrumental variables results presented below for the child health response variables. Virtually identical results are obtained when we consider the first-stage reduced forms that correspond to the expenditures per adult equivalent, school enrollment or fertility response variables.

All specification include a rich set of covariates x_{ihc}: We include child sex, age (in months) and age squared, log household size, 8 dummies for the educational attainment of the household head, 12 dummies for the type of water source used by the household (for the child anthropometric results), 27 dummies for the household head's language (which largely controls for ethnic-specific heterogeneity), 6 provincial dummies, as well as the distance to Luanda and to the provincial capital. Standard errors are clustered at the *comuna* level in order to account for common shocks affecting all observations within a given *comuna* .

Note that we could have attempted to measure the long-run effect of the Angolan conflict by restricting our sample to households living in areas that were only affected by conflict in the 1975-1994 period (and not in the 1995-2000 period), and the short-run effect by considering a subsample of individuals living in areas that only experienced conflict during the 1995-2000 period (with no conflict during the 1975-1994 period). We do not do so, for two reasons. First, all areas that were subject to conflict during the 1995-2000 period were also affected by conflict during the earlier period –there is thus no way of isolating the effect due solely to relatively recent episodes of conflict. Second, if we restrict our

[18] Results are qualitatively similar when we replace total casualties with either military or civilian casualties. The same is true of the structural relationships presented in what follows. In the interests of brevity, we restrict our attention to the total casualty measures of conflict intensity. Full results are available upon request.

attention to areas that only suffered from conflict during the 1975-1994 period (and were not affected by conflict in the earlier period) or sample becomes extremely small. Moreover, since the theoretical motivation that lies behind our identification strategy (the effect of the distance to rebel headquarters on conflict intensity) has little to say concerning the dynamics of conflict in a given area over time, we prefer to restrict our attention to two relatively simple questions: (i) do conflict episodes that occurred between 1975 and 1994 have an impact on, for example, the anthropometrics of children born *after* the violence in question obtained? And what is the impact of conflict episodes that occurred during the 1995-2000 period, which corresponds in most cases to children born during the episodes of violence in question? The answer to the first question is our measure of the long-run effects of conflict, the answer to the second, our measure of its short-run impact. Though limited in scope, our questions have the merit of being answerable relatively rigorously, given the data and the identification strategy at hand.

Answering these two questions will allow us in what follows to show three things. First, there is a long-run effect of conflict, which is persistent, and which affects individuals who were not directly affected by the conflict *per se*. Second, this long-run effect appears to depreciate rapidly over time. Third, the short-run effects on individuals directly exposed to civil war can be disastrous, which should not come as much of a surprize.

As should be clear from the coefficients reported, for example, in column 1 —where the dependent variable is the total number of conflict-related deaths during the 1975-1994 period, within a 1 km radius of the community— the decreasing and convex relationship predicted by our theoretical model of civil war obtains, with the distance from Jamba reducing the total number of casualties, at a decreasing rate. Evaluated at the median sample distance, the marginal effect of an additional kilometre from Jamba is to reduce total casualties by 3.4 individuals. The same qualitative results emerge for the 5, 10 and 20 km radii. The coefficients reported in the right-hand portion of the Table, which correspond to the 1995-2000 period are smaller (though just as statistically significant), given that the average number of casualties during this period is significantly smaller. Again, the decreasing and convex relationship between conflict intensity and distance to UNITA headquarters predicted by our model obtains. Note also that the distance separating the community from Luanda is associated with (statistically significant) lower levels of conflict intensity, as is the distance from the corresponding provincial capital.

The negative effect of the distance to rebel headquarters mirrors that found by Hegre and Raleigh (2007) in their study of the determinants of conflict events in 14 Central African countries using the Armed Conflict Location and Events Dataset (ACLED). In contrast, they find that the likelihood of conflict events increases with distance to the capital.[19] Hagelstein (2008) uses distance variables such as those considered here in order to explain the geographical pattern of conflict events during the Algerian civil war.

1.6 Empirical results

It will be of some interest for the empirical results that follow to ascertain whether the observed effects of civil war on our outcome variables can be given a coherent theoretical interpretation. In order to do so, and in order to be able to formulate testable hypotheses, we begin by constructing a unitary Neoclassical household model that links child nutritional status, school enrollment, household expenditures and fertility behavior.

1.6.1 Baseline linear instrumental variables results

Baseline linear instrumental variables results are presented in Tables 1.5 and 1.6 for child anthropometrics, in Table 1.7 for log household expenditures per adult equivalent, in Table 1.8 for school enrollment, and in Table 1.9 for the total number of live births.

As shown in the right-hand portion of Tables 1.5 and 1.6, the marginal impact of conflict intensity during the 1995-2000 period on child health in 2000, be this measured by height-

[19] Buhaug (2007), using a specification in which the dependent variable is the distance from the capital to the conflict event, finds that relatively capable insurgents tend to operate relatively close to the capital.

for-age or weight-for-age $z-$scores, is negative and statistically significant for the 1, 5 and 10 km radii. It becomes indistinguishable from zero once the 20 km radius is reached. In quantitative terms, the point estimates are large: each additional 1,000 conflict-related deaths between 1995 and 2000 within a 1 km radius reduces the WAZ $z-$score in 2000 by 3.518, with a standard error of 0.886. If we compare a community with zero conflict-related deaths within 1 km to a community with 44 deaths (the sample standard deviation), the difference in the corresponding WAZ $z-$score would be $0.044 \times 3.518 = 0.154$. This is 13% of the sample standard deviation in WAZ $z-$scores (1.176). Moving from a situation in which there are no casualties within 1km to one in which there are 104 (the sample maximum), decreases the WAZ $z-$score in 2000 by $0.104 \times 3.518 = 0.365$.

These results contrast sharply with the impact of casualties suffered between the onset of the civil war and the Lusaka Protocols (the 1975-1994 period). For the 1km radius, each additional $1,000$ casualties reduces the WAZ $z-$score by 0.279 ($se = 0.065$), a marginal effect that is more than 12 times smaller than the effect of casualties during the 1995-2000 period. The corresponding numbers are similar for the HAZ $z-$scores.

In the lower portion of Tables 1.5 and 1.6, we report the corresponding OLS estimates obtained *without* instrumenting conflict intensity. For the child anthropometrics response variables, the raw OLS coefficient (with no covariates) associated with conflict intensity is *positive* and statistically significant for all radii, when we consider casualties suffered during the 1975-1994 period. Once covariates x_{ihc} are included, however, the point estimates become statistically indistinguishable from zero. Taken together, these results are in conformity with what we would expect if unobservables that increase conflict intensity also improve child health, as argued above in section 1.4, and as spelled out explicitly in the expression for the likely bias of OLS-based estimates given in equation (1.6). Note also that the null hypothesis of the exogeneity of conflict intensity is soundly rejected for all radii, for both child anthropometric response variables. For the 1995-2000 measure of conflict intensity, the differences between the IV and OLS results are less striking, though exogeneity is strongly rejected and the OLS point estimates are all positive, though only statistically significant for the 20km radius.

The results reported in Table 1.7 show that casualties suffered during the 1995-2000 period reduce log household expenditures per capita, though the effect is measured very imprecisely (the coefficient is equal to -1.333 for the 1km radius, with a standard error of 1.446). Exogeneity is rejected in all cases. In contrast, casualties suffered during the 1975-1994 period yield a very small increase in log household expenditures per capita (between 3.5 and 7.2%) for the 5 to 20 km radii (with the effect being statistically indistinguishable from zero for the 1km radius), and exogeneity is *not* rejected: if one were therefore to base oneself in this case on the OLS point estimates with covariates (the lowermost line of results reported in the Table) one would conclude that conflict intensity during the 1975-1994 period has no statistically significant impact on household expenditures per adult equivalent in 2000.

Table 1.8 presents results for school enrollment in 2000 where, for children between 5 and 12 years of age, the dependent variable is equal to one if the child is in school at the time of the survey, and zero otherwise. The regression should therefore be viewed as a linear probability model, and the coefficients interpreted as approximations to the marginal effects at the sample mean. For the 1, 2 and 5km radii, the results presented in Table 1.8 indicate that the likelihood of a child being in school is *increasing* in conflict intensity during the 1995-2000 period. For the 1km radius, for example, an increase of one standard deviation in casualties (0.044) increases the probability that a child is in school by $0.044 \times 2.402 = 0.105$. As with household expenditures per adult equivalent, the results for conflict intensity during the 1975-1994 period are much smaller (and statistically indistinguishable from zero, except for the 1km radius). Again, as with the household expenditure results, the exogeneity of conflict intensity is rejected for casualties during the 1995-2000 period, but not (apart from the 2km radius) for the 1975-1994 period.

Table 1.9 presents results for fertility behavior (the total number of live births per woman of reproductive age). Conflict intensity during the 1995-2000 period significantly decreases fertility for the 1 and 2km radii: a one standard deviation increase in conflict intensity for the 1km radius leads to a reduction of $0.044 \times 2.108 = 0.092$ live births per woman. In contrast, conflict intensity during the 1975-1994 period has no statistically significant effect on fertility.[20] Exogeneity is not rejected for either period.

Taken as a whole, the findings reported in Tables 1.5 to 1.9 are what would be predicted by the Neoclassical household model presented above, with perhaps the most striking result being the short-term *positive* effect of conflict intensity on school enrollment, as in the model whose comparative statics results were presented in column 1 of Table 1.4. Indeed, given that the comparative statics of school enrollment in this case are given by $dt^* = -\frac{T\alpha\beta}{(w-\alpha)^2(\beta+\gamma-\delta+\epsilon)}dw > 0$, one would expect the impact of civil war to be particularly large when the expected future returns to investment in educational human capital, parameterized by α, are substantial. The short-term effect of conflict intensity on log household expenditures per adult equivalent, $dc^* = \frac{\theta}{\theta+\delta}dy$, which is statistically indistinguishable from zero, is consistent with households with a very low elasticity of consumption (the parameter θ in the utility function $U(i, c, t, x, n) = i^\epsilon c^\theta t^\beta x^\gamma n^\delta$) and hence a high degree of risk-aversion, and a relatively high preference for large family size (high δ). Both of these interpretations square well with the Angolan context, as does the extremely important short-term impact of conflict intensity on child anthropometrics, $dx^* = \frac{T\gamma}{p(\beta+\gamma-\delta+\epsilon)}dw$, which is consistent with households putting a great deal of weight on child nutritional status (large γ). Finally, the negative short-term impact of conflict intensity on live births presented in Table 1.9 implies that the negative impact on fertility of the fall in exogenous household income outweighs the positive impact of falling wage rates.

[20] Results are similar if we replace linear IV with a Poisson model.

1.6.2 Selection issues due to infant survival and displaced persons

There are two issues that could bias our results due to the selection problems that they might induce. The first, which is potentially important when it comes to measures of nutritional outcomes, is child mortality. As should be obvious, for a given level of conflict intensity it is likely that it is the healthier children who survive. The greater the level of conflict, the more severe this selection issue is likely to be. Though one could in theory correct for this using a standard Heckman procedure (the identification for the first stage probit could be achieved using the same exclusion restriction as for conflict intensity, as first introduced into the literature by Thomas and Strauss (1997), we have no information in our dataset concerning individuals who died during the conflict. It is likely that this form of selection would tend to *reduce* the absolute value of our estimate of the deleterious effect of conflict intensity on child health.

The second problem that could potentially bias our results stems from displaced persons. In essence, displaced persons generate a discrepancy between the geographical area where the conflict occurred and the geographical area in which the individuals considered by the survey were subjected to conflict. If the individuals being surveyed arrived relatively recently in the area, it could be that they lived most of their lives in an area whose level of conflict intensity has little to do with that of area where they now reside. A priori, one would expect individuals to have moved from areas where conflict intensity was relatively high to areas where conflict intensity is relatively low. This would lead to an understatement (in absolute value terms) of the impact of conflict intensity on observed outcomes In order to ascertain the potential magnitude of this effect, Tables 1.10, 1.11, 1.12, 1.13 and 1.14 present all of our specifications, but where the sample has been restricted to individuals born within their province of residence: this should minimize or at least reduce the understatement of the impact of conflict due to displaced persons. As should be clear from the reported coefficients, almost all of our results are both qualitatively and, reassuringly, quantitatively, preserved. Tests of the difference in key parameter estimates between the two subsamples indicate that, for the school enrollment, log expenditures per adult equivalent and number of live births response variables, none are significant. For the other outcomes (the anthropometric response variables) the differences are statistically significant and, as such, we have included the displacement dummy variable in all specifications presented in Tables 1.10, 1.11 so as to account for this difference.

1.6.3 The time profile of war's impact

A particularly striking aspect of the empirical results presented so far is the contrast between short- and long-term effects of conflict intensity on our response variables. For child anthropometrics, school enrollment and fertility, the short-term effects were all statistically significant and large, with the corresponding long-term impacts being either small or statistically indistinguishable from zero. Though unsurprizing in light of the existing empirical evidence summarized in section 1.2, and which shows that violent conflict has little if any long-run effects, the unfortunate length of the Angolan conflict (27 years overall, 25 years of which correspond to our data) provides one with an interesting natural experiment within which to attempt to quantify the approximate rate at which the effect of conflict "depreciates" or indeed reverses itself.

In what follows, we divide our response variables into two categories: (i) child anthropometrics, for which the long- and short-run effects will turn out to be of the same sign when both pre-1994 and post-1995 casualties are included simultaneously in the structural equation; (ii) household expenditures per adult equivalent, school enrollment and the total number of live births, where the opposite pattern will be found.

Consider first child anthropometrics. In Table 1.15, where our response variables are weight-for-age and height-for-age $z-$scores in 2000, we include *both* pre-1994 and 1995-2000 casualties, for the 1km radius. Let β_1 represent the coefficient associated with casualties in a given pre-1994 period, and β_2 the coefficient associated with casualties in the 1995-2000 period. Then if the marginal impact of casualties on the response variable is constant, one can infer the "depreciation rate" θ associated with casualties in the earlier period by solving $\beta_1 = \beta_2 \exp\{-\theta k\}$, where k represents the amount of time separating the most recent from the earlier period. This yields:

$$\theta = \frac{1}{k} \ln\left(\frac{\beta_2}{\beta_1}\right).$$

For the HAZ, and for the results presented in column (1), for example, this implies that the approximate annual rate of depreciation of casualties is equal to 63.1%.

As an aggregation exercise (and drawing inspiration from the vintage capital literature, see, e.g. Benhabib and Rustichini, 1991), we construct discounted stocks of casualties, expressed in 1995-2000 casualty equivalents, using each of the 4 five-year pre-1994 periods. Figure 1.3 plots the cumulative distribution functions of casualties (in thousands) for the 1975-79, 1980-84, 1985-89, 1990-1994 and 1995-2000 periods. The right-hand lowermost plot represents the cumulative distribution function of discounted casualties, where exponential discounting based on the results from column (1) is applied.[21] Alternatively, we can apply period-specific discount rates given by the ratios of the coefficients associated with the most recent and earlier periods in each of the four columns.[22] This leads to more weight being placed on earlier periods in that the ratios of coefficients are all roughly similar. Empirical results using both definitions of discounted casualties are presented for both response variables in Table 1.15. As expected, and given the extremely high rate at which effective casualties depreciate (implying that most of the variation in conflict intensity will be due to the most recent period), the point estimates are all extremely similar to those reported earlier (see Table 1.6) for the 1995-2000 period.

Consider now our second category of response variables (log household expenditures per adult equivalent, school enrollment and the total number of live births). In Table 1.16, we

[21] More explicitly, if x_t represents casualties in period t, it corresponds to $x = x_{1995-2000} + \frac{\beta_1}{\beta_2}\left(x_{1990-94} + \frac{x_{1985-89}}{2} + \frac{x_{1980-84}}{3} + \frac{x_{1975-79}}{4}\right)$.

[22] The corresponding formula is $x = x_{1995-2000} + \frac{\beta_{11}}{\beta_{21}} x_{1990-94} + \frac{\beta_{12}}{\beta_{22}} x_{1985-89} + \frac{\beta_{13}}{\beta_{23}} x_{1980-84} + \frac{\beta_{14}}{\beta_{24}} x_{1975-79}$, where the second subscript on the coefficients refers to the corresponding column in the Table 1.15.

present the same specification as in Table 1.15, in that we include *both* pre-1994 and 1995-2000 casualties, for the 1km radius. For household expenditures, the impact of pre-1994 casualties is now larger than in the results presented in Table 1.7, positive, and statistically significant at the usual levels of confidence. Conversely, the impact of casualties during the 1995-2000 period is negative and, in contrast to the results presented in Table 1.7, is measured with much greater precision. The same pattern emerges for the number of live births, with a positive and statistically significant effect of pre-1994 casualties, and a negative and statistically significant effect of casualties during the 1995-2000 period.

For fertility, this pattern is consistent with the available empirical evidence from *developed* countries (summarized earlier in section 1.2) which suggests that a fall in fertility during periods of conflict is followed by a rebound in post-conflict periods. The empirical results presented in Table 1.16 partially disentangle these two effects despite the entire 1975-2000 period being characterized by conflict, and show that, given the appropriate data and a sufficiently long time span (25 years in this case), the hypothesized pattern does indeed obtain. Much the same can be said for the results for household expenditures: civil war leads to a sharp drop in household expenditures in the short-run, with a catching up phenomenon –consistent with convergence in a Vintage capital growth model, in which more recent vintages of physical and human capital are more productive— in the long-run. For schooling, the large, positive and statistically significant short-run effect of casualties in the 1995-2000 period remains; pre-1994 casualties decrease the net enrollment ratio, but this effect is not statistically significant.

1.6.4 Cross-sectional nonlinearities

While the impact of conflict intensity on our response variables follows a non-linear pattern *over time*, with (i) large short-run effects rapidly disappearing so that no quantitatively important effects remain after 5 years for child anthropometrics, and (ii) short-run and long-run effects of opposite signs emerging for household expenditures and fertility, non-linearities may also exist at the cross-sectional level, with different marginal impacts obtaining at different *levels* of conflict intensity. In order to explore such cross-sectional non-linearities, we consider a simple semiparametric specification of the form:

$$y_{ihc} = x_{ihc}\alpha + f(w_c) + \varepsilon_{ihc},$$

where w_c is allowed to have a completely general relationship with our different response variables. A nontrivial complication that arises in this context is that standard nonparametric techniques are not applicable, in that, more often than not, w_c is correlated with ε_{ihc}.

A number of nonparametric two-stage least squares estimators are available. Here, we adopt that proposed by Newey and Powell (2003), with minor modifications. In particular, we replace their specification, which is based on a Hermite series approximation and a researcher-specified constraint parameter for their functional compactness restriction, with a specification based on cubic thin plate splines with the smoothing parameter chosen automatically by restricted maximum likelihood (REML).[23] Montecarlo simulation results based on the same experimental design as in their paper, and presented in Appendix A, show that our estimator has finite sample properties that are almost identical to theirs. Selected empirical results, corresponding to cases in which particularly interesting non-linearities were uncovered, are presented in the four panels of Table 1.17.

In order to facilitate direct comparison with our linear IV results, the graphs reported in Table 1.17 plot the first derivative $\frac{\partial y_{ihc}}{\partial w_c} = \frac{\partial f(w_c)}{\partial w_c}$ of the response variable with respect to conflict intensity; each figure also includes a horizontal dotted line that plots the corresponding (constant) marginal effect from linear IV reported earlier in the appropriate column of Tables 1.5 to 1.9.[24] Given the two-stage nature of the nonparametric 2SLS

[23] The functional compactness restriction is necessary because nonparametric 2SLS yields an "ill posed inverse" problem in the mapping between the reduced form and structural parameters. On the estimation of the smoothing parameter by REML, see e.g. Ruppert, Wand, and Carroll (2000).

[24] Note that derivative estimation is an ongoing area of research in nonparametric statistics. For ex-

procedure, the horizontal axis plots the level of casualties *predicted* by our nonparametric first-stage reduced forms.

For the 1km radius WAZ $z-$score results, Panel A of Table 1.17 shows that conflict begins to have a statistically significant negative impact only once the 120 casualty level is reached (i.e. once the upper, +2 standard deviation, confidence band lies *below* the horizontal axis) and is measured less and less precisely as the number of casualties increases. For the corresponding 5 km radius, presented in Panel B, things are even more complex: a substantial negative marginal impact obtains for very low levels of conflict intensity, then vanishes, and finally reappears between 80 and 150 casualties. It is clear in this case that the linear IV estimates —represented in the graph by the horzontal dotted line (and which were presented in Table 1.5) are almost entirely driven by this threshold effect that appears once the 80 casualty level for the 1995-2000 period is considered. Results for the HAZ $z-$scores were qualitively similar (not reported in the interest of brevity).

Results for fertility behavior (the total number of live births) for the 10 km radius are presented in Panel C of Table 1.17. As with the WAZ $z-$scores for the 1 km radius, the marginal impact of conflict intensity is statistically indistinguishable from zero until roughly the 120 casualty level is reached (and is measured too imprecisely once one goes beyond the 160 casualty level to be statistically significant). Finally, Panel D presents the results for log household expenditures per adult equivalent for the 10 km radius. Here, the pattern is similar to that found for the WAZ $z-$scores for the 5 km radius (and presented in Panel B): very substantial and statistically signficant marginal effects are found at very low levels of conflict intensity, followed by an interval in which the marginal impact is statistically indistinguishable from zero, followed by a final interval, roughly similar to that found for the WAZ $z-$scores for the 5 km radius, in which the marginal impact of casualties is large, negative, statistically significant, and measured with an acceptable degree of precision.

ample, Jarrow, Ruppert, and Yu (2004) note that: "there is a general lack of robust derivative estimators. Smoothing parameter selection tools are based on optimising the estimate of the regression function and not of the derivative, which can lead to serious undersmoothing." A good primer on the topic is provided by Newell and Einbeck (2007). Results in levels are, of course, available upon request.

Though it is difficult to establish a definitive taxonomy of cross-sectional non-linearities, the results presented in Table 1.17 are suggestive of two patterns. First, for response variables such as child anthropometrics and household expenditures, coupled with conflict intensity measures at the 5 and 10 km radii respectively, there are substantial negative marginal impacts at very low levels of conflict intensity, which then vanish, only to reappear once a threshold (of 80 casualties) is crossed. Second, for child anthropometrics and a very small radius, and fertility and a broader radius, no statistically significant effects of conflict appear at very low levels of conflict intensity, while once a higher threshold (of roughly 120 casualties for the 1995-2000 period) is crossed, the marginal impact of conflict becomes negative, large and statistically significant, though measured with growing imprecision.

Though, as with all nonparametric procedures, our results depend upon the degree of smoothing that is imposed, our modification of the Newey and Powell estimator to allow a "pluggin method" (REML) to be applied removes part of the discretion afforded to the researcher, and affords some degree of objectivity. Moreover, our findings suggest that there are significant cross-sectional non-linearities in the impact of conflict on household outcomes that, to the best of our knowledge, have heretofore been ignored by researchers in both the economics and political geography literatures, and which warrant additional investigation.

1.7 Concluding remarks

This chapter has considered the impact of 25 years of the Angolan civil war on child anthropometrics, household expenditures per adult equivalent, school enrollment and fertility. Our identification strategy was based on the plausibly exogenous variation in conflict intensity generated by the distance separating each community in our sample from the logistical base of UNITA, the principal rebel group. For historical reasons, and in order to study differences between the long- and short-term effects of conflict on household outcomes, we distinguished between two periods of the civil war: (i) 1975-1994, which allows us to measure the long-term effects, and (ii) 1995-2000, which allows us to focus on short-term effects.

Linear instrumental variables estimates, based on household survey data collected in 2000, indicate that, in the short-run, conflict intensity worsens child health, does not significantly affect household expenditures, increases school enrollment and decreases fertility, as would be predicted by a Neoclassical unitary household model inspired by the work of Rosenzweig (1990). In contrast, in the long-term (for casualties suffered before the Lusaka Protocols) we find a small though statistically significant impact of conflict intensity on child health, and no impact on household expenditures, enrollment and fertility.

While our short-term results follow the pattern predicted by our unitary household model, our long-term results do not. In part, this is probably due to the "depreciation" over time of the marginal impact of conflict intensity. Our long-term results indicate that, even in the absence of an explicit reconstruction policy in areas that suffered particularly heavily during the conflict (as was the case, for example, in Vietnam, see Miguel and Roland, 2005), communities can recover relatively quickly —and this in a context in which violence was ongoing. Inspired by the vintage capital literature, we tested this hypothesis explicitly by simultaneously including short- and long-term effects in a single equation. For child anthropometrics, relatively small or statistically insignificant negative long-term effects coexist with much larger negative short-term effects, implying an extremely steep "depreciation rate" for "effective" casualties. For household expenditures per adult equivalent and fertility, on the other hand, small *positive* long-term effects coexist with large negative short-term effects.

The pattern associated with fertility is similar to that observed in developed countries during and after conflicts, and it is perhaps only the unfortunate length of the Angolan conflict that allows us to identify it, in contrast to most available evidence for conflicts in developing countries. The pattern associated with household expenditures, for its part, suggests vintage capital effects in the context of a Neoclassical growth model, in which conflict-related destruction of physical and human capital yields a large short-term drop in expenditures, followed by recovery to a higher steady-state level of expenditures fuelled by the greater productivity of more recent vintages.

Though the negative long-term effects of conflict intensity on child anthropometrics are quantitatively small, they are —at least for the case of height-for-age— statistically significant. Thus, children born, for example, in 1999, still suffer from a statistically significant deficit in height if they lived in communities which suffered from particularly high levels of violence up to 24 years earlier. This result, as well as the disastrous *short-run* effect of conflict intensity on child anthropometrics suggest clear policy priorities — emergency nutrition programs being the most obvious— both *during* conflicts and in the post-conflict recovery period. This is brought home by available empirical evidence that indicates that poor early childhood health is associated with substantially less schooling during adolescence, worse adult health, and lower adult productivity (Appleton, 2001, Alderman, Hoddinott, and Kinsey, 2006). Finally, the mechanics of our household model, combined with our empirical results, suggests that it may be worthwhile to devote some effort to collecting information on the evolution of prices and wages during civil wars in order to more clearly isolate the specific channels through which household behavior is affected by violent conflict.

1.8 Annexe 1 : Montecarlo evidence on our implementation of the Newey and Powell (2003) nonparametric 2SLS estimator

In this appendix, we provide Montecarlo simulation evidence on our implementation of the Newey and Powell (2003) nonparametric 2SLS estimator. The experimental design is identical to theirs (see pp. 1574-5), with the structural equation being given by:

$$y = g(x) + u = \ln(|x-1|+1)\mathrm{sgn}(x-1) + u,$$

the first stage reduced form by:

$$x = z + v,$$

and where the disturbance terms u and v and the instrument z are generated as:

$$\begin{pmatrix} u \\ v \\ z \end{pmatrix} \sim \text{i.i.d.}N\left(0, \begin{pmatrix} 1 & 0.5 & 0 \\ 0.5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right).$$

For a sample size of 400 observations, the root mean-squared-error (RMSE), averaged across the realized values of x and the 500 replications yields a value of 0.233, which lies between the values (0.208 and 0.356) presented by Newey and Powell for specifications based on two different values of the constraint parameter for their functional compactness restriction. Figure 1.2 provides a visual representation of the results of our Montecarlo simulation. The right-hand panel represents the results of nonparametric 2SLS estimation, with the thick black line being the estimate and the dotted lines the ± 2 standard error confidence interval. The thinner grey line represents the "true" relationship. The straight grey line represent the linear IV estimate which fails to account for the nonlinearity of the

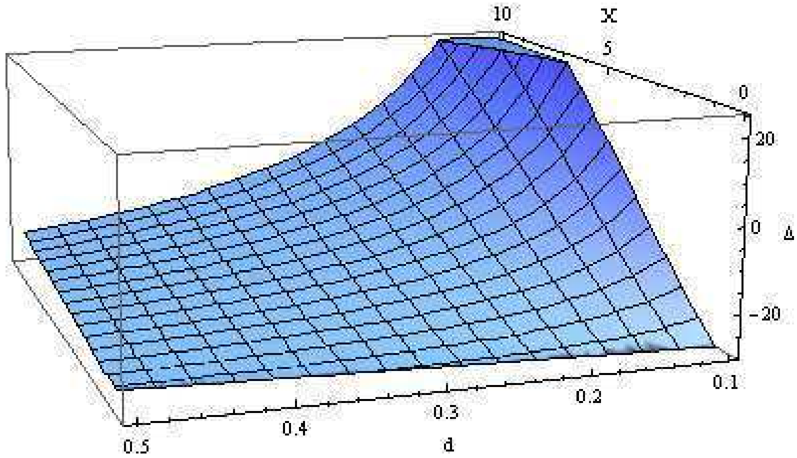


Figure 1.1: A rent-seeking model of civil war. $\Delta = \Pi^W - \Pi^D$ (likelihood of conflict) on the vertical axis as a function of distance from rebel headquarters d, and size of the prize X. Civil war breaks out when $\Delta > 0$, otherwise deterrence prevails.

structural relationship. The left-hand panel represents a naive nonparametric estimate of the structural relationship which fails to correct for endogeneity, with the straight line representing the OLS estimate. It is obvious from the left-hand figure that failing to correct for endogeneity often results in extremely precise —and incorrect— inference: the true relationship lies, more often than not, outside the ± 2 standard error confidence interval. This is not the case with the nonparametric 2SLS estimates, which always include the true relationship within the ± 2 standard error confidence interval.

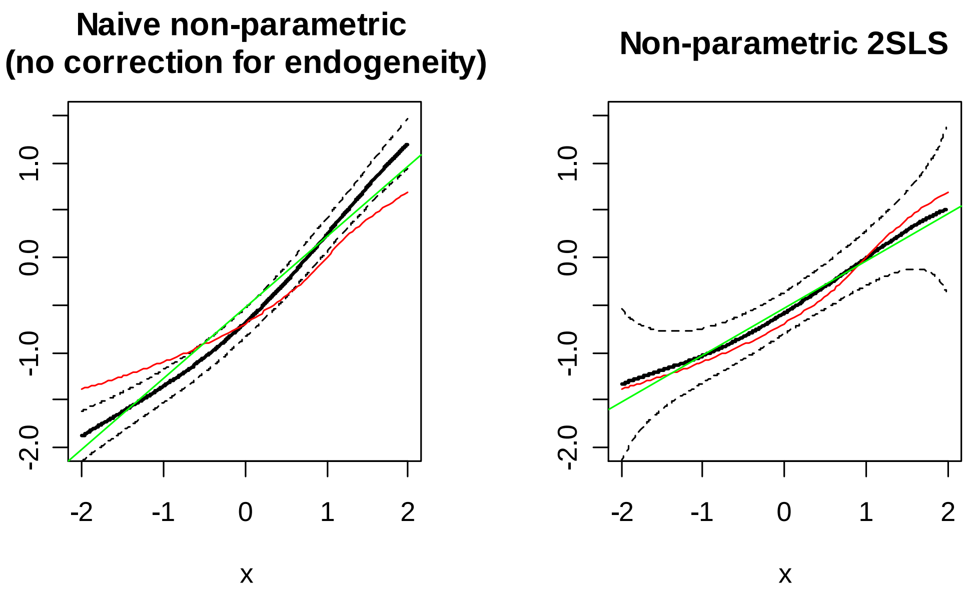


Figure 1.2: Thin plate spline implementation of the Newey and Powell (2003) nonparametric 2SLS estimator. Smoothing parameter chosen automatically by restricted maximum likelihood (REML). True relationship ($y = \ln(|x-1|+1)\text{sgn}(x-1)+u$) given by thin grey line. Thick black line represents the nonparametric estimate (with ± 2 standard error confidence bands). Montecarlo simulation results with 400 observations and 500 replications.

	mean	median	std.
Response variables: y_{ihc} or y_{hc}			
Child height-for-age (HAZ) $z-$score	−1.577	−1.620	1.433
Child weight-for-age (WAZ) $z-$score	−1.203	−1.260	1.176
Household expenditures			
per adult equivalent (Kwanzas)	390	249	477
Child is enrolled in school	0.507	1	0.499
Total number of live births	4.05	3.00	2.81
Covariates: x_{ihc}, x_{hc} and x_c			
Child age (months)	28.09	28.00	17.10
Child is female	0.490	0.000	0.499
Child is Kimbundo	0.469	0.000	0.499
Years education of household head	4.537	5.000	2.056
Acces to water in the house	0.120	0.000	0.325
Household size	7.27	7.00	3.17
Distance to:			
Luanda, in 10 km	38	38	37
Provincial capital, in 10 km	8	3	13
Exclusion restrictions: z_c			
Distance to UNITA headquarters, in 10 km			
Jamba	116	114	24

Table 1.1: Descriptive statistics: response variables and covariates (8328 observations for child-specific response variables, 5247 observations for household-specific variables).

	Conflict intensity: w_c 1975-1994			Conflict intensity: w_c 1995-2000		
	mean	median	std	mean	median	std
Total casualties (thousands) within:						
1 km	0.907	0.580	1.788	0.016	0.000	0.044
5 km	0.844	0.026	1.632	0.020	0.000	0.048
10 km	1.120	0.073	2.007	0.038	0.000	0.059
20 km	1.794	0.726	2.452	0.102	0.025	0.140
Civilian casualties (thousands) within:						
1 km	0.498	0.316	0.883	0.007	0.000	0.044
5 km	0.416	0.176	0.777	0.010	0.000	0.048
10 km	0.606	0.316	0.992	0.025	0.000	0.059
20 km	0.573	0.358	0.895	0.056	0.014	0.140
Military casualties (thousands) within:						
1 km	0.549	0.173	1.014	0.009	0.000	0.031
5 km	0.512	0.021	0.963	0.009	0.000	0.031
10 km	0.489	0.023	0.963	0.013	0.000	0.031
20 km	0.983	0.398	1.307	0.045	0.002	0.091

Table 1.2: Descriptive statistics: conflict intensity variables, 1975-1994 and 1995-2000.

	Dependent variable: total casualties (w_c) in thousands							
	1975-1994, within:				1995-2000, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
Exclusion restrictions: z_c								
Distance to UNITA headquarters, in 10 km								
Jamba	−0.490	−0.580	−1.133	−0.783	−0.048	−0.048	−0.060	0.040
	(0.038)	(0.044)	(0.033)	(0.022)	(0.000)	(0.001)	(0.000)	(0.002)
Jamba, sq	0.002	0.003	0.006	0.004	0.000	0.000	0.000	−0.000
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Covariates x_{ihc}								
Distance from								
Luanda	−0.051	−0.104	−0.130	−0.101	−0.001	−0.001	−0.001	−0.008
	(0.004)	(0.005)	(0.003)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
Prov. cap.	−0.039	−0.095	−0.094	−0.090	0.000	0.000	0.000	−0.013
	(0.006)	(0.007)	(0.005)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)
$\overline{R}^2$	0.204	0.293	0.809	0.913	0.58	0.54	0.86	0.77
F^{62}_{8265}	35	56	572	1417	190	164	866	474
[$p-$value]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Weak IV diagnostics: partial								
R^2	0.024	0.041	0.195	0.210	0.271	0.227	0.484	0.136
$F-$ statistic	3.31	4.05	5.18	2.68	21.36	21.74	26.58	2.42
[$p-$value]	[0.045]	[0.023]	[0.009]	[0.078]	[0.000]	[0.000]	[0.000]	[0.099]

Table 1.3: First-stage reduced forms of the determinants of conflict intensity, 8,328 observations. Child-, household- and community-specific covariates included in all specifications. Standard errors clustered at the comuna level in parentheses.

	Human capital production function		
	$h(t,x)=t$	$h(t,x)=x$	$h(t,x)=t+x$
Wage rate			
$\frac{dx^*}{dw}$	$\frac{T\gamma}{p(\beta+\gamma-\delta+\epsilon)}>0$	$\frac{T\gamma}{(p-\alpha)(\beta+\gamma-\delta+\epsilon)}>0$	$\frac{T\gamma}{(p-\alpha)(\beta+\gamma-\delta+\epsilon)}>0$
$\frac{dt^*}{dw}$	$-\frac{T\alpha\beta}{(w-\alpha)^2(\beta+\gamma-\delta+\epsilon)}<0$	0	$-\frac{T\alpha\beta}{(w-\alpha)^2(\beta+\gamma-\delta+\epsilon)}<0$
$\frac{dn^*}{dw}$	$-\frac{y(\beta+\gamma-\delta+\epsilon)}{Tw^2(\delta+\theta)}<0$	$-\frac{y(\beta+\gamma-\delta+\epsilon)}{Tw^2(\delta+\theta)}<0$	$-\frac{y(\beta+\gamma-\delta+\epsilon)}{Tw^2(\delta+\theta)}<0$
$\frac{dc^*}{dw}$	0	0	0
Exogenous household income			
$\frac{dx^*}{dy}$	0	0	0
$\frac{dt^*}{dy}$	0	0	0
$\frac{dn^*}{dy}$	$\frac{\beta+\gamma-\delta+\epsilon}{Tw(\delta+\theta)}>0$	$\frac{\beta+\gamma-\delta+\epsilon}{Tw(\delta+\theta)}>0$	$\frac{\beta+\gamma-\delta+\epsilon}{Tw(\delta+\theta)}>0$
$\frac{dc^*}{dy}$	$\frac{\theta}{\delta+\theta}>0$	$\frac{p(2\epsilon+\theta)-\alpha(2(\gamma+\epsilon)+\theta)}{(p-\alpha)(\delta+\theta)}\lesseqgtr 0$	$\frac{\theta}{\delta+\theta}>0$
Price of food			
$\frac{dx^*}{dp}$	$-\frac{Tw\gamma}{p^2(\beta+\gamma-\delta+\epsilon)}<0$	$-\frac{Tw\gamma}{(p-\alpha)^2(\beta+\gamma-\delta+\epsilon)}<0$	$-\frac{Tw\gamma}{(p-\alpha)^2(\beta+\gamma-\delta+\epsilon)}<0$
$\frac{dt^*}{dp}$	0	0	0
$\frac{dn^*}{dp}$	0	0	0
$\frac{dc^*}{dp}$	0	$\frac{2y\alpha\gamma}{(p-\alpha)^2(\delta+\theta)}>0$	0
Overall impact of civil war: $dw<0, dy<0, dp>0$			
dx^*	$\frac{T\gamma}{p(\beta+\gamma-\delta+\epsilon)}dw - \frac{Tw\gamma}{p^2(\beta+\gamma-\delta+\epsilon)}dp<0$	$\frac{T\gamma}{(p-\alpha)(\beta+\gamma-\delta+\epsilon)}dw - \frac{Tw\gamma}{p^2(\beta+\gamma-\delta+\epsilon)}dp<0$	$\frac{T\gamma}{(p-\alpha)(\beta+\gamma-\delta+\epsilon)}dw - \frac{Tw\gamma}{p^2(\beta+\gamma-\delta+\epsilon)}dp<0$
dt^*	$-\frac{T\alpha\beta}{(w-\alpha)^2(\beta+\gamma-\delta+\epsilon)}dw>0$	0	$-\frac{T\alpha\beta}{(w-\alpha)^2(\beta+\gamma-\delta+\epsilon)}dw>0$
dn^*	$-\frac{y(\beta+\gamma-\delta+\epsilon)}{Tw^2(\delta+\theta)}dw + \frac{\beta+\gamma-\delta+\epsilon}{Tw(\delta+\theta)}dy \lesseqgtr 0$	$-\frac{y(\beta+\gamma-\delta+\epsilon)}{Tw^2(\delta+\theta)}dw + \frac{\beta+\gamma-\delta+\epsilon}{Tw(\delta+\theta)}dy \lesseqgtr 0$	$-\frac{y(\beta+\gamma-\delta+\epsilon)}{Tw^2(\delta+\theta)}dw + \frac{\beta+\gamma-\delta+\epsilon}{Tw(\delta+\theta)}dy \lesseqgtr 0$
dc^*	$\frac{\theta}{\delta+\theta}dy<0$	$\frac{p(2\epsilon+\theta)-\alpha(2(\gamma+\epsilon)+\theta)}{(p-\alpha)(\delta+\theta)}dy + \frac{2y\alpha\gamma}{(p-\alpha)^2(\delta+\theta)}dp \lesseqgtr 0$	$\frac{\theta}{\delta+\theta}dy<0$

Table 1.4: Comparative statics of nutrition, educational attainment, fertility and household consumption per capita.

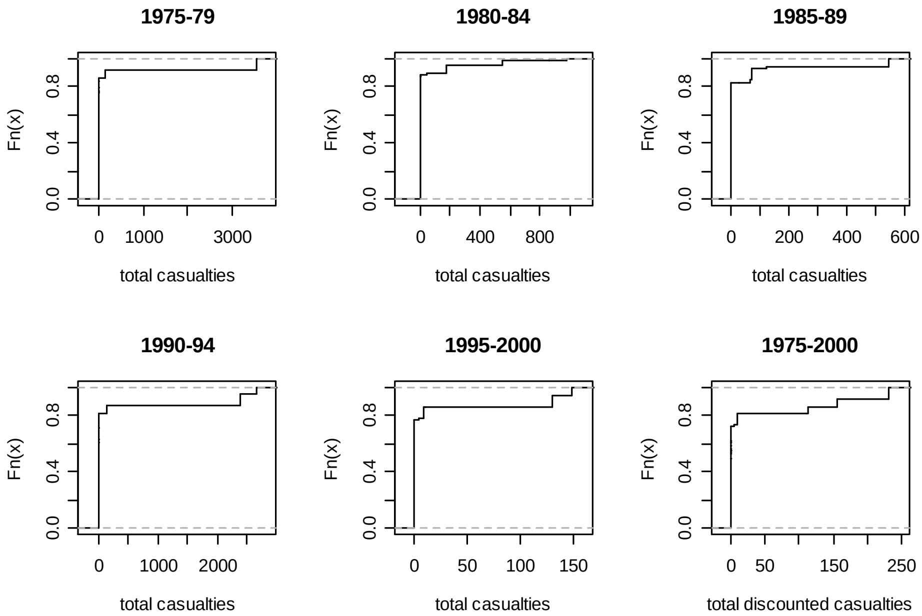


Figure 1.3: Cumulative distributions of total casualties and total discounted casualties.

Dependent variable y_{ihc} : child weight-for-age (WAZ) $z-$score in 2000								
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
β_{IV}	-0.279 (0.065)	-0.138 (0.041)	-0.087 (0.023)	-0.115 (0.032)	-3.518 (0.886)	-3.540 (0.890)	-2.956 (0.702)	1.215 (0.851)
Test of exogeneity of w_c : $p-$value								
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.031
β_{OLS}, covariates:								
not included	0.035 (0.017)	0.025 (0.016)	0.024 (0.012)	0.028 (0.010)	0.024 (1.062)	0.209 (0.973)	0.521 (0.714)	1.005 (0.164)
included	0.016 (0.012)	0.014 (0.011)	0.005 (0.018)	0.009 (0.017)	-0.184 (0.481)	-0.177 (0.427)	-1.004 (0.574)	0.288 (0.211)

Table 1.5: Instrumental variables estimates of the impact of conflict intensity on child weight-for-age (WAZ) z-score, 8,328 observations, child- and household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{ihc} : child height-for-age (HAZ) $z-$score in 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
β_{IV}	-0.345 (0.080)	-0.187 (0.051)	-0.114 (0.028)	-0.153 (0.039)	-3.781 (0.956)	-3.810 (0.949)	-3.289 (0.719)	1.788 (1.487)
Test of exogeneity of w_c : $p-$value								
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
β_{OLS}, covariates:								
not included	0.034 (0.014)	0.033 (0.013)	0.029 (0.012)	0.031 (0.010)	0.365 (0.738)	0.687 (0.759)	0.912 (0.616)	0.965 (0.148)
included	0.001 (0.011)	0.015 (0.011)	0.011 (0.022)	-0.028 (0.024)	-0.837 (0.466)	-0.389 (0.485)	-1.095 (0.858)	0.118 (0.298)

Table 1.6: Instrumental variables estimates of the impact of conflict intensity on child height-for-age (HAZ) z-score, 8,328 observations, child- and household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{hc} : log household expenditures per adult equivalent in 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
β_{IV}	0.036 (0.033)	0.072 (0.023)	0.035 (0.013)	0.047 (0.017)	-1.333 (1.446)	-1.305 (1.463)	-0.672 (1.291)	-1.836 (1.573)
Test of exogeneity of w_c : $p-$value								
	0.892	0.491	0.524	0.674	0.001	0.006	0.006	0.000
β_{OLS}, covariates:								
not included	0.098 (0.040)	0.082 (0.045)	0.097 (0.043)	0.119 (0.040)	1.550 (2.216)	2.316 (2.186)	3.634 (2.048)	2.652 (0.806)
included	0.011 (0.012)	0.012 (0.011)	-0.009 (0.015)	-0.008 (0.022)	-0.170 (0.888)	-0.301 (0.727)	-1.323 (0.817)	-0.469 (0.192)

Table 1.7: Instrumental variables estimates of the impact of conflict intensity on log household expenditures per adult equivalent, 10,014 observations, household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{hc} : child is enrolled in school in 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
β_{IV}	0.020	0.006	0.004	0.005	2.402	2.418	1.776	0.133
	(0.010)	(0.007)	(0.004)	(0.005)	(0.709)	(0.731)	(0.735)	(0.664)
Test of exogeneity of w_c : $p-$value								
	0.224	0.016	0.942	0.742	0.000	0.000	0.000	0.000
β_{OLS}, covariates:								
not included	0.015	0.022	0.019	0.019	0.869	0.469	1.010	0.315
	(0.011)	(0.010)	(0.006)	(0.007)	(0.314)	(0.345)	(0.287)	(0.136)
included	0.002	0.011	0.007	-0.002	0.725	0.295	1.185	0.062
	(0.012)	(0.010)	(0.012)	(0.010)	(0.697)	(0.323)	(0.624)	(0.191)

Table 1.8: Instrumental variables estimates of the impact of conflict intensity on school enrollment. 14,361 observations, household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{hc} : total number of live births, up until 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1km	5km	10km	20km	1km	5km	10km	20km
β_{IV}	-0.044	0.017	0.004	0.009	-2.108	-2.136	-1.358	-0.922
	(0.085)	(0.060)	(0.033)	(0.044)	(1.079)	(1.094)	(0.894)	(1.190)
Test of exogeneity of w_c : $p-$value								
	0.841	0.505	0.865	0.597	0.518	0.167	0.823	
β_{OLS}, covariates:								
not included	-0.077	-0.027	-0.006	-0.013	-3.486	-1.485	-1.833	-1.742
	(0.035)	(0.035)	(0.027)	(0.023)	(0.913)	(0..902)	(1.318)	(0.547)
included	-0.064	-0.029	-0.001	0.034	-2.675	-0.744	-1.231	-0.254
	(0.027)	(0.023)	(0.023)	(0.030)	(0.760)	(0.399)	(0.740)	(0..319)

Table 1.9: Instrumental variables estimates of the impact of conflict intensity on number of live births, 10,365 observations, household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{ihc} : child weight-for-age (WAZ) $z-$score in 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
β_{IV}	-0.278 (0.066)	-0.130 (0.047)	-0.085 (0.026)	-0.112 (0.037)	-3.819 (0.822)	-3.834 (0.829)	-3.140 (0.690)	1.091 (0.774)
Test of exogeneity of w_c : $p-$value								
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.055
β_{OLS}, covariates:								
not included	0.037 (0.016)	0.029 (0.015)	0.026 (0.012)	0.029 (0.009)	0.076 (1.158)	0.309 (1.051)	0.591 (0.724)	0.985 (0.160)
included	0.016 (0.011)	0.017 (0.010)	0.010 (0.017)	-0.003 (0.016)	-0.307 (0.549)	-0.219 (0.470)	-1.029 (0.603)	0.260 (0.203)

Table 1.10: Instrumental variables estimates of the impact of conflict intensity on child weight-for-age (WAZ) z-score, 7,886 observations. Sample restricted to individuals born within their province of residence. Child- and household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{ihc} : child height-for-age (HAZ) $z-$score in 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
β_{IV}	-0.348 (0.070)	-0.182 (0.054)	-0.114 (0.029)	-0.152 (0.042)	-4.150 (0.907)	-4.175 (0.900)	-3.531 (0.690)	1.682 (1.379)
Test of exogeneity of w_c : $p-$value								
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
β_{OLS}, covariates:								
not included	0.037 (0.013)	0.038 (0.013)	0.033 (0.012)	0.033 (0.010)	0.425 (0.815)	0.799 (0.825)	1.032 (0.624)	0.932 (0.140)
included	0.002 (0.011)	0.018 (0.011)	0.016 (0.022)	-0.023 (0.024)	-0.928 (0.501)	-0.425 (0.524)	-1.117 (0.897)	0.103 (0.290)

Table 1.11: Instrumental variables estimates of the impact of conflict intensity on child height-for-age (HAZ) z-score, 7,886 observations. Sample restricted to individuals born within their province of residence. Child- and household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{hc} : log household expenditures per adult equivalent in 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
β_{IV}	0.062 (0.220)	0.163 (0.108)	0.115 (0.081)	0.134 (0.094)	−1.005 (1.797)	−0.945 (1.817)	−0.517 (1.789)	−1.029 (0.457)
Test of exogeneity of w_c : $p-$value								
	0.801	0.192	0.221	0.302	0.946	0.858	0.007	0.100
β_{OLS}, covariates:								
not included	0.090 (0.068)	0.061 (0.077)	0.103 (0.061)	0.149 (0.055)	0.495 (2.646)	1.538 (2.962)	3.338 (2.909)	2.304 (0.909)
included	−0.001 (0.015)	0.013 (0.018)	0.000 (0.024)	0.001 (0.035)	−0.959 (0.973)	−1.066 (0.883)	−1.982 (0.911)	−0.737 (0.179)

Table 1.12: Instrumental variables estimates of the impact of conflict intensity on log household expenditures per adult equivalent, 5,023 observations. Sample restricted to individuals born within their province of residence. Household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{hc} : child is enrolled in school in 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1 km	5 km	10 km	20 km	1 km	5 km	10 km	20 km
β_{IV}	0.022 (0.020)	0.008 (0.015)	0.005 (0.008)	0.006 (0.010)	0.565 (0.709)	0.557 (0.336)	0.371 (0.227)	0.115 (0.314)
Test of exogeneity of w_c : $p-$value								
	0.228	0.533	0.922	0.799	0.000	0.000	0.167	0.081
β_{OLS}, covariates:								
not included	0.022 (0.011)	0.029 (0.009)	0.025 (0.007)	0.025 (0.007)	0.974 (0.333)	1.208 (0.354)	1.177 (0.295)	0.341 (0.134)
included	−0.001 (0.002)	−0.000 (0.002)	0.005 (0.003)	0.004 (0.003)	−0.145 (0.123)	−0.026 (0.143)	0.226 (0.173)	−0.098 (0.063)

Table 1.13: Instrumental variables estimates of the impact of conflict intensity on school enrollment. 12,068 observations. Sample restricted to individuals born within their province of residence. Household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Dependent variable y_{hc} : total number of live births, up until 2000							
	1975-1994				1995-2000			
	Total casualties (w_c), in thousands, within:				Total casualties (w_c), in thousands, within:			
	1km	5km	10km	20km	1km	5km	10km	20km
β_{IV}	-0.262 (0.194)	-0.049 (0.103)	-0.039 (0.066)	-0.034 (0.078)	-3.380 (1.493)	-3.343 (1.496)	-2.779 (1.155)	-0.217 (0.461)
Test of exogeneity of w_c : $p-$value	0.293	0.866	0.432	0.195	0.697	0.645	0.464	0.941
β_{OLS}, covariates:								
not included	-0.176 (0.034)	-0.107 (0.072)	-0.155 (0.031)	-0.162 (0.025)	-4.197 (1.649)	-4.436 (1.784)	-6.505 (1.638)	-1.084 (0.604)
included	-0.055 (0.033)	-0.032 (0.024)	0.009 (0.030)	0.063 (0.018)	-2.975 (0.928)	-2.854 (0.825)	-2.223 (1.114)	-0.246 (0.289)

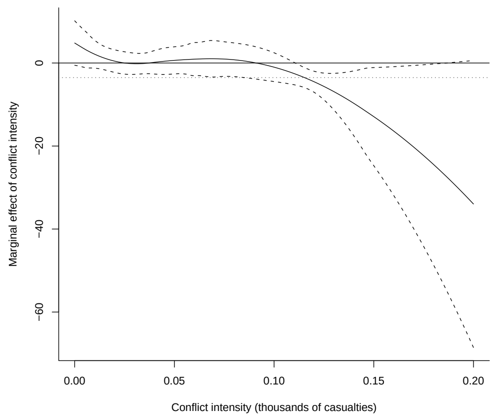
Table 1.14: Instrumental variables estimates of the impact of conflict intensity on number of live births, 5,904 observations. Sample restricted to individuals born within their province of residence. Household-specific covariates, distance to Luanda and to provincial capital, and province dummies included. Standard errors clustered at the comuna level in parentheses.

	Initial year for pre-1994 period			
	1990	1985	1980	1975
	Weight-for-age $z-$score in 2000			
Total casualties: pre-1994	-0.067 (0.067)	-0.071 (0.071)	-0.073 (0.073)	-0.048 (0.047)
Total casualties: 1995-2000	-3.346 (0.606)	-3.102 (0.753)	-3.149 (0.720)	-3.129 (0.727)
First-stage reduced forms:				
$F-$statistics	82.03, 165.5	82.4, 165.5	83.3, 165.5	56.6, 165.5
R^2s	0.380, 0.553	0.382, 0.553	0.384, 0.553	0.298, 0.553
Total discounted casualties:	$\sum_{t=0}^{t=4} x_t e^{-\delta t}$		$\sum_{t=0}^{t=4} x_t e^{-\delta_t}$	
1975-2000	-3.345 (0.476)		-3.028 (0.434)	
First-stage reduced forms:				
$F-$statistics, R^2s	155, 0.463		99.68, 0.427	
	Height-for-age $z-$score in 2000			
Total casualties: pre-1994	-0.144 (0.059)	-0.152 (0.061)	-0.156 (0.063)	-0.103 (0.048)
Total casualties: 1995-2000	-3.392 (0.652)	-2.872 (0.735)	-2.972 (0.718)	-2.929 (0.779)
First-stage reduced forms:				
$F-$statistics	82.0, 165.5	82.4, 165.5	83.3, 165.5	56.6, 165.5
R^2s	0.380, 0.553	0.382, 0.553	0.384, 0.553	0.298, 0.553
Total discounted casualties:	$\sum_{t=0}^{t=4} x_t e^{-\delta t}$		$\sum_{t=0}^{t=4} x_t e^{-\delta_t}$	
1975-2000	-3.380 (0.531)		-2.769 (0.441)	
First-stage reduced forms:				
$F-$statistics, R^2s	92.2, 0.409		78.9, 0.371	

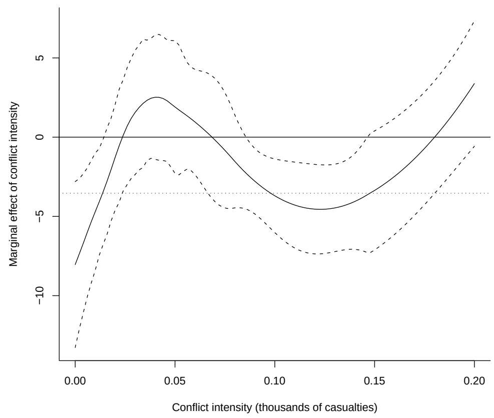
Table 1.15: Inferring the "depreciation rate" on casualties. Standard errors clustered at the comuna level in parentheses.

	Initial year for pre-1994 period			
	1990	1985	1980	1975
	Log household expenditures per adult equivalent in 2000			
Total casualties: pre-1994	0.263 (0.069)	0.278 (0.073)	0.287 (0.073)	0.187 (0.050)
Total casualties: 1995-2000	−2.532 (1.214)	−3.469 (1.358)	−3.331 (1.226)	−3.433 (1.169)
First-stage reduced forms:				
$F-$statistics	64.41, 139.6	64.92, 139.6	65.38, 139.6	56.6, 165.5
R^2s	0.389, 0.58	0.391, 0.58	0.392, 0.58	0.298, 0.553
	Child is enrolled in school in 2000			
Total casualties: pre-1994	−0.026 (0.026)	−0.028 (0.028)	−0.029 (0.029)	−0.018 (0.019)
Total casualties: 1995-2000	0.690 (0.266)	0.785 (0.321)	0.784 (0.323)	0.787 (0.333)
First-stage reduced forms:				
$F-$statistics	90.11, 170.2	90.07, 170.2	90.82, 170.2	61.28, 170.2
R^2s	0.388, 0.545	0.388, 0.545	0.390, 0.545	0.301, 0.545
	Total number of live births, up until 2000			
Total casualties: pre-1994	0.217 (0.111)	0.231 (0.118)	0.241 (0.123)	0.151 (0.078)
Total casualties: 1995-2000	−2.984 (1.252)	−3.759 (1.469)	−3.648 (1.432)	−3.697 (1.482)
First-stage reduced forms:				
$F-$statistics	94.13, 218.3	95, 218.3	95.65, 218.3	68.14, 218.3
R^2s	0.372, 0.599	0.374, 0.579	0.376, 0.579	0.308, 0.579

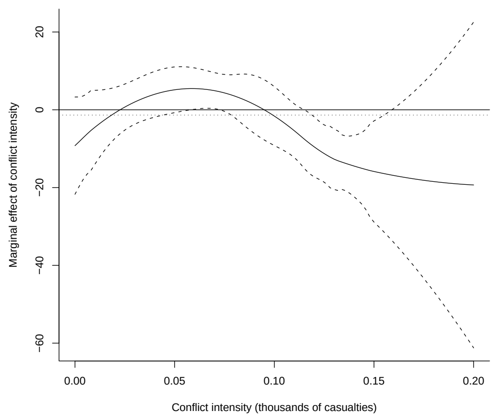
Table 1.16: The time profile of the impact of casualties: log household expenditures per adult equivalent, school enrollment, and fertility. Standard errors clustered at the comuna level in parentheses.



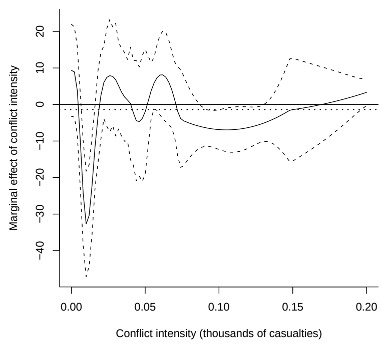
Panel A: WAZ,
1 km radius



Panel B: WAZ,
5 km radius



Panel C: Fertility,
10 km radius



Panel D: Log expenditures
per adult equivalent, 10 km radius

Table 1.17: Selected semiparametric Newey-Powell 2SLS estimates of the marginal impact of conflict intensity (dotted curves are plus/minus 2 standard error confidence bands). Horizontal dotted line plots the corresponding marginal effect estimated by linear IV.

Chapter 2

How effective are social programs during conflicts? Evidence from the Angolan civil war

2.1 Introduction

Post-conflict countries face huge problems. These include weak state capacity, the destruction of physical and human infrastructure, widespread poverty and massive unemployment. It is thus crucial, during the post-conflict recovery process, to have some idea of what type of programs should be prioritized so as to speed up recovery and reduce the likelihood that the country will fall once again into the conflict trap. This is particularly important in that authors such as Collier (2007) have argued that post-conflict countries have a 50% likelihood of relapsing into civil war during the first decade of peace, and evidence presented by Miguel, Satyanath, and Sergenti (2004) suggests that income shocks and poverty are the main driving force behind civil wars in subsaharan Africa.

One of the key instrument used by the World Bank to promote the poverty reduction in developing countries today is constituted by Community Driven Development (CDD) programs and Social Funds. The CDD approach is used by the World Bank across a broad range of countries in support of a variety of urgent needs: water supply and sewage rehabilitation; school and health center construction, nutrition programs for mothers and infants; the building of rural access roads and support for micro-enterprises. Currently, the World Bank's CDD portfolio amounts to approximately $ 2 billion a year (Mansuri and Vijayendra 2004).

While a majority of CDD programs and social funds implemented since the first Bolivian social fund of 1987 have been so in countries at peace, the approach has also been applied in conflict or post-conflict countries. Early examples include the 1994 Angola social fund (FAS I) and the 1998 Indonesia Kecamatan Development Program. According to one accounting(Owen and Bannon 2006), 94 CDD projects were in operation in 2006 in countries characterized by violent armed conflict or post-conflict reconstruction.. To the best of our knowledge, none of these programs have been the subject of a properly carried out impact evaluation. For two reasons, the Angolan context provides a unique opportunity to assess whether social programs of the CDD/social fund variety can be effective. First, because of the length of the Angolan civil war –27 years, and second because it was one of the first countries to implement a social fund in a conflict situation.

The two questions we attempt to answer in this chapter are very much policy-driven. First, how effective are social programs during conflicts? Second, how is the effectiveness of social programs affected by conflict intensity? Answering these two questions allows one to ascertain whether social programs can be an effective anti-poverty tool in conflict situations, and thus mitigate impact of civil war on human capital and whether a particular targeting mechanism —in which treatment is prioritized on the basis of the conflict intensity suffered by a given geographical unit— yields a higher marginal benefit from treatment. If such a targeting mechanism is not implemented, the consequence is that the overall benefits of the program could have been significantly higher, without additional expenditure being needed.

In what follows, we consider the impact on child anthropometrics and household expenditures of the only major social program in Angola during the civil war —the first two phases of the World Bank sponsored *Fundo de Apoio Social* (FAS I and II). Our identification strategy is based on the political geography of program deployment. Our hypothesis is that one of the purposes of program implementation was to consolidate government political support in areas contested with UNITA, the main rebel group, as well as to maintain the support of the population in areas to the rear of frontline communities. Based on a simple model of spatial competition in the tradition of Hotelling (1929), we show that the likelihood of FAS deployment should be an increasing function of the distance separating a given community from the government's forward bases in Bailundo (Huambo province), and a decreasing function of the distance to UNITA's main base at the time in Andulo (Bie province). Using these exclusion restrictions to generate plausibly exogenous variation in treatment status, and unique household data collected in Angola in 2000, our linear instrumental variables estimates show that treatment by the FAS during the 1994-2000 period was associated with a 48.5% (se = 0.175) increase in household expenditures per adult equivalent, and a 0.359 ($se = 0.166$) increase in child height-for-age $z-$scores. We then use the local instrumental variables estimator to explore the marginal treatment effect (MTE) of the FAS. In particular, we consider the interaction between treatment by the FAS and the intensity of conflict, using the remarkable data on episodes of violence during the Angolan civil war compiled by Ziemke (2007). In doing so, we find that each 1,000 additional deaths attributable to the civil war within a 20 km radius of the community shifts the MTE associated with household income per adult equivalent up by 18.2%. The FAS was therefore associated with substantial benefits for treated communities, with these benefits being significantly increasing in the intensity of violence faced by the inhabitants.

The remainder of this chapter is organized as follows. Section 2.2 describes the socio economic situation in Angola after several years of war. Section 2.3 describes the functioning of the FAS, which is typical of World Bank-funded social funds operating in various parts of the world. It also summarizes the existing evidence on the impact of social funds, which is largely limited to their effects on educational outcomes, with relatively little attention having been devoted to household expenditures and child anthropometrics. Section 2.4 presents the motivation of the choices of our outcomes indicators while section 2.5 explains the program deployment in conflit situation.

Section 2.6 presents our identification strategy and develops a theoretical model of the spatial division of political support between the government and UNITA, the main rebel group in Angola during the period of the civil war under consideration. Section 2.7 presents our basic empirical specification, and demonstrates how, given targeting on communities that have unobservables that make them relatively poor, OLS-based estimates of the impact of the FAS on child anthropometrics and household expenditures are likely to be biased downwards, therefore calling for instrumental variables-based methods. We also present the household survey data, collected in 2000, on which our estimates will be based, and report the results of the first-stage reduced forms for treatment status. The latter do not reject the predictions of our theoretical model and appear to provide an appropriate degree of identification.

In section 2.8 we present our empirical results. We divide these into four parts. Section 2.8.1 presents linear instrumental variables (IV) results that show that treatment by the FAS was associated with highly significant and quantitatively large increases in household expenditures per adult equivalent and child height-for-age $z-$scores. The corresponding OLS-based estimates are negative and often statistically significant (depending upon whether covariates are included or not), highlighting the importance of adequately controlling for the endogeneity of treatment status. Section 2.8.2 explores heterogeneity in program impact along different demographic and location characteristics. Section 2.8.3 presents the possible transmission channels of FAS on our outcomes indicators and section 2.8.4 tests for the presence of essential heterogeneity using the linearity test described (among other sources) in Heckman, Urzua, and Vytlacil (2006). We reject the absence of essential heterogeneity for household expenditures per adult equivalent, and do not for child height-for-age $z-$scores. Closer scrutiny of the household expenditures results using a carefully specified Roy model reveals a great deal of heterogeneity in the impact of treatment by the FAS, with an inverted $U-$shaped pattern emerging for the marginal treatment effect: the effect of treatment was negative (though marginally significant) for communities with unobservables that made them either highly likely or highly unlikely to receive treatment while, for communities with unobservables in between, the impact of treatment was positive and highly significant.

Section 2.9 combines our household survey data with extremely detailed conflict intensity data collated by Ziemke (2007). Continuing with the Roy model-based analysis of section 2.8.4, we show that, in communities which were subjected to particularly high levels of conflict, the MTE of the FAS was particularly large: an increase of 1,000 casualties within 20km of the community over the 1994-2000 period is associated with an increase of 18.2% ($se = 0.026$) in the marginal treatment effect of the FAS on household expenditures per adult equivalent. Section 2.10 concludes, and provides some tentative thoughts on the implications of our results for the deployment of social programs in situations of conflict.

2.2 The context

Following independance in 1975, Angola went through a 27 year civil war that only came to an end in 2002 with the physical elimination of UNITA leader Jonas Savimbi. Despite (or perhaps because of) important petroleum resources and diamonds, as well as a large endowment of productive arable land, GDP per capita in 1994 was estimated at $US 420, and much of the country's significant pre-independance infrastructure had been destroyed. Between 500,000 and 1m persons had been killed, and 9m landmines peppered the countryside. A household survey conducted in 1994 by the national statistical agence (INE) estimated that 64% of the population was living below the poverty line, whereas another study produced by Fernando Ribeiro in 1993 put the figure at 86%. Infant mortality for 0-5 year olds was estimated to be a staggering 320 per 1,000 live births –perhaps not surprizing in that the INE estimated in 1995 that most health centers in rural areas of the country had been destroyed.

Much the same could be said of the educational system, with many schools either having been destroyed or closed. Student-to-teacher ratios skyrocketed, with average class size in Luanda running at 140 in the early 1990s. Primary enrollment during this period was estimated at 50%, with secondary enrollment at 12%. During the same period, it was estimated that only 30% of the population had access to potable water (with access being better in urban areas): this was reflected in statistics concerning water-borne diseases, with frequent cases of cholera outbreaks and high prevalence rates of infantile diarrhea. Food insecurity was estimated to affect 70% of the population displaced during the civil war. Overall, agricultural productivity fell substantially: from an average output of 100kg of wheat equivalent per capita in 1980, to 67kg in September 1993. In nutritional terms, this is equivalent to a dietary energy supply (DES) per capita of 600 kcal/day, with the benchmark minimum for the Angolan context being estimated at 2,100 kcal/day.

2.3 The FAS and the impact of social funds

The FAS is an autonomous structure created by the Angolan government in October 1994. It is placed under the authority of the Ministry of Planning (MINPLAN) which determines its objectives. Between 1994 and 2000, the first two phases of the FAS (I and II) were deployed in 9 out of the 18 Angolan provinces: Cabinda, Luanda, Kwanza Sul, Benguela, Huambo, Namibe, Huila, Cunene and Bengo. The areas of intervention of the FAS I and II are shown, along with the date of program deployment, on the map of Angola displayed in Figure 2.2. By and large, program deployment was confined to areas under government control, though certain provinces, such as Huambo, had only recently been retaken from UNITA, the main rebel group headed by Jonas Savimbi. An important purpose of the FAS was to aid in the transition towards peace (which, unfortunately, only lasted until 1996, when fighting broke out again until the final conclusion of the conflict in 2002), by targeting households that had been affected by the conflict and the ensuing economic crisis. The stated aims of the program were to improve community access to basic infrastructure, to improve the capacity of communities and local NGOs to plan, evaluate, manage and maintain community-level infrastructure, to create income generating activities both in rural and urban areas, and to contribute to an understanding of the sources of poverty in order to formulate more effective poverty reduction strategies.

The functioning of the FAS is similar to that of many social funds around the world. Local NGOs (*agencias de enquadramento*) work alongside communities (represented by the *nucleo comunitario*) so as to identify priority infrastructures and to prepare projects which are then submitted to the FAS. Provincial FAS offices vet submitted project proposals in coordination with the provincial government. All approved projects receive FAS funding, and their implementation is carried out by the communities themselves who must contribute 10% of total project costs, either in kind, through labor, or in cash. Between 1994 and 2001, the FAS funded $29m worth of projects; 67% of these funds were provided by multilateral donors such as the World Bank, with the rest coming from bilateral donors and the Government of Angola (GOA).[1] A total of 685 projects were funded during the period, with 34% in education, 9% in health, 47.4% in water and waste management, and 9% in income generating activities. The average size of a FAS project is $20,000.[2]

There is a growing literature on the impact of social funds though, with one exception, very little is known concerning their effects in an African context. Four papers, published in a special issue of the *World Bank Economic Review*, applied pipeline and propensity score matching methods (as well, to a more limited extent, difference-in-differences (DID), experimental and instrumental variables approaches) to evaluate the impact of social funds in Bolivia, Nicaragua, Armenia and Peru.

Evia, Newman, Pradhan, Ramiro, Rawling, and Ridder (2002) consider the impact of small-scale health, water and educational rural infrastructure funded by the Bolivian Social Investment Fund (SIF) over the 1993-1994 period. On the basis of experimental methods, they find that SIF-funded educational projects signficantly improved the quality of school infrastructure (as measured by the number of books per student, class size, the presence of a latrine, clean drinking water and electricity in the school), while there was no impact on school enrollment rates, absenteeism, and test scores. The infrastructure did nevertheless significantly reduce dropout rates in primary classes by 3.8%. Propensity score matching- and DID-based evidence indicates that SIF-funded projects in health and water infrastructure increased usage and significantly reduced infant mortality rates.

[1] We thank Victor Hugo Guilherme and Henda Ducados, respectively Director and Deputy-Director of the FAS III, for these numbers, which they presented at the World Bank Conference on Community-Driven Development, Ouagadougou, Burkina Faso, November 3-4, 2003.

[2] These numbers come from the FAS administrative database. The level of World Bank funding was far below that settled on in the IDA credit agreement with the Angolan government, which amounted to $54m. The total level of FAS I and II funding is relatively small when compared with social funds in Latin America. Disbursements of the Bolivian SIF between 1994 and 1998 amounted to $160m, those of the Nicaraguan FISE between 1991 and 1998 amounted to $191m while disbursements of the Peruvian FONCODES between 1992 and 1998 totaled $472m.

Pradhan and Rawlings (2002) consider the impact of social infrastructure funded by the Nicaragua Emergency Social Investment Fund (FISE) between 1991 and 1998. They use pipeline and propensity score matching and show, based on household data collected in 1998, that educational projects significantly increased school enrollment, reduced absenteism, and reduced the age of enrollment in the first grade. Health infrastructure investments were associated with greater usage of health centers, though there were no significant impacts on child anthropometrics. Similarly, FISE investments in water and waste management infrastructure had no discernible impact on child anthropometric outcomes.

Chase (2002) uses pipeline and propensity score matching methods to analyse educational and water infrastructure financed by the Armenian Social Investment Fund (ASIF) during the 1996-2000 period. Educational investments increased household expenditures on educational items, and increased enrollment rates. Water projects significantly increased access to potable water.

In contrast to the three preceding papers, where the results are almost always based on the assumption of selection on observables, Paxson and Schady (2002) consider the impact of the Peruvian Social Fund (FONCODES) over the 1993-1996 period using instrumental variables techniques. Using a community-level poverty index and the proportion of citizens having voted against the previous regime during the 1990-1993 period as exclusion restrictions, they show that FONCODES educational projects increased school enrollment rates for children between 6 and 11 years of age.

A fifth paper, also based on pipeline and propensity score matching methods, provides evidence on a social fund in Zambia (the MPU), a neighbor of Angola, which may be the most relevant in our case. Chase and Sherbume-Benz (2001) show that educational projects significantly increased school attendance, increased household educational expenses, whereas health projects, though they increased usage as well as vaccination rates, did not significantly affect health outcomes such as child diarhea.

2.4 The Outcomes and treatement

In what follows we will consider two outcome indicators, the height for age $z-$score (using the standard WHO/CDC methodology) of children of between 0 and 5 years of age, and expenses per adult equivalent. The first indicator is widely used to assess the impact of access to clean water and healthcare on child health (Block and Webb (2003); David, Moncada, and Ordonez (2004); Christiaensen and Alderman (2004); Valdivia (2004); Galiani, Gertler, and Schargrodsky (2005); Leipziger, Fay, Wodon, and Yepes (2005)). It is a long-run indicator that should reflect chronic malnutrition and repeated spells of disease, and which is not sensitive to short-run shocks. The second indicator is measured over a one month reference period. We prefer expenditures to revenues because of the usual arguments linked to the cyclical nature of household revenues (particular in rural areas) and the ability, even in a poor country context, of households to smooth their consumption patterns over time. As such, expenditures per adult equivalent may be an acceptable proxy for a household's permanent income.

Though the FAS did devote 9% of its funds to income generating activities, the bulk of the program went to infrastructures such as schools, health centers, transportation infrastructure (such as bridges) and clean water. Examples of previous analyses of anti-poverty programs that have used expenditures per adult equivalent as their outcome measure include Ravallion (1991, 1999), Besley and Coate (1992), and Lipton and Ravallion (1995).

Note that treatment status is defined slightly differently for our two outcome variables. For child anthropometrics, a community (*comuna*) is assumed to be treated if it received FAS expenditures in health, water, or waste disposal. For expenditures per adult equivalent, we add educational and income generating project activities.

2.5 The program deployment in conflit situation

A fundamental problem faced by program initiators in conflict situations is the informational basis upon which program deployment is determined. While reasonably accurate poverty maps, upon which targeting decision could be based, are often a rarity in poor countries, they are literally non-existent in conflict situations. The same can be said of usable household surveys, especially in the context of a civil war as long as Angola's.

The *analysis* of how CDD or social funds are targeted is usually carried out at two levels. Poverty targeting involves analyzing the extent to which resources are channeled to the poorer segments of the population. In turn, the analysis of poverty targeting is usually carried out a two levels of aggregation –geographic targeting and household targeting. Working at lower levels of disaggregation can often be important. For instance, a simple per capita income decomposition undertaken by Paxson and Schady (2002) using the 1998 Peru LSMS showed that only 22% of the underlying variance was due to between-community differences. The second level of analysis is preference targeting, in which the focus is on whether the types of implemented subprojects match the *ex-ante* preferences expressed by the potential beneficiaries –this form of analysis is seldom carried out when it comes to CDD or social funds, perhaps because such programs are by their very nature believed to be "bottom up".

In practical terms, and thus in terms of deployment, much of the targeting for social funds and CDD programs obtains through the process of community bidding. Community bidding, in turn, is fettered by the criteria used by program initiator to score projects, such as explicit pro-poor preferences or positive discrimination in favor of projects that are intended to benefit ethnic minorities or individuals with disabilities. Available evidence (on the Nicaragua FISE, Zambia ZamSIF, Peru FONCODES, and Armenia SIF) yields mixed findings concering the effectiveness of poverty targeting in CDD programs. While community-level targeting would appear to be progressive, household-level targeting is much less successful, leading to the well-known "elite capture" critique of CDD programs. Moreover, as noted in a widely publicized report by the World Bank on CDD (World Bank, 2004), results often differ between rural and urban areas. In countries in the midst of conflict, such as Angola, it is likely that program deployment is driven by concerns that may have little to do with poverty alleviation. Indeed, conversations with individuals who were associated with FAS program deployment during its inception leads us to believe that, logically enough, the program was at least in part seen as a manner of consolidating support of communities for the government. In the section that follows, we use this logic in order to motivate our identification strategy.

2.6 Identification strategy

2.6.1 Intuition

The identification strategy that we adopt in this chapter is based on the political geography underlying the deployment of the first two phases of the FAS. Our excluded instruments are given by the distance separating a given community (i) from the seat of the *União Nacional para a Independência Total de Angola* (UNITA) headquarters at the time of program initiation, and (ii) from the forward government bases which had just been captured in the context of an offensive in late 1994. More specifically, we will focus on the distance to the towns of Andulo (in Bie province) which served as UNITA military headquarters following the signing of the Lusaka Protocol in 1994 (see Spears 1999), which brought a temporary lull in the conflict and Bailundo (in Huambo province), a former UNITA base that constituted the forward prong of the government's advance into the central highlands.[3]

Given that it is likely that one of the purposes of the government, through program deployment, was to maximize the political benefit from doing so, it is highly probable that

[3]Other important UNITA headquarters were located and Lumbala N'guimbo in Moxico province, near the western border of Zambia, and at Jamba, in the southeastern corner of Cuando Cubango province, just north of the slice of Namibia known as the Caprivi strip. See Angola Peace Monitor (2000).

the geographic location of a community with respect to Andulo and Bailundo was one of the factors determining the decision on where to deploy the FAS program. In particular, our hypothesis is that program deployment in areas located between government forward bases in Huambo province and UNITA headquarters in Andulo was geared towards gaining political support in these areas, while simultaneously maintaining support in the rear. We now turn to a spatial model that will help us to sharpen this intuition and provide a theoretical basis for our first-stage reduced forms.

2.6.2 A spatial model of treatment status

Consider a spatial competition model in the tradition of Hotelling (1929).[4] The government and the UNITA rebels are competing for political support in the area lying between the forward bases of the government in Bailundo (Huambo province), which was conquered in the 1994 offensive, and Andulo (Bie province), to which UNITA withdrew prior to the ceasefire brought about by the Lusaka Protocols. We assume that potential supporters are distributed uniformly along the line linking Bailundo to Andulo. The purpose of the government, through the geographical allocation of FAS program expenditures, is to maximize its level of support on the ground between a given community and Andulo, net of the cost of providing the program. Moreover, the government wishes to deploy the program in such a manner that it maintains support in areas "behind" the community; this includes areas located between Bailundo and the community, as well as areas located further back, and which are not directly contested by UNITA.

[4] A more formal illustration of our identification strategy would be based on a model of spatial competition with two characteristics (in this case, longitude and latitude), such as Economides (1986).

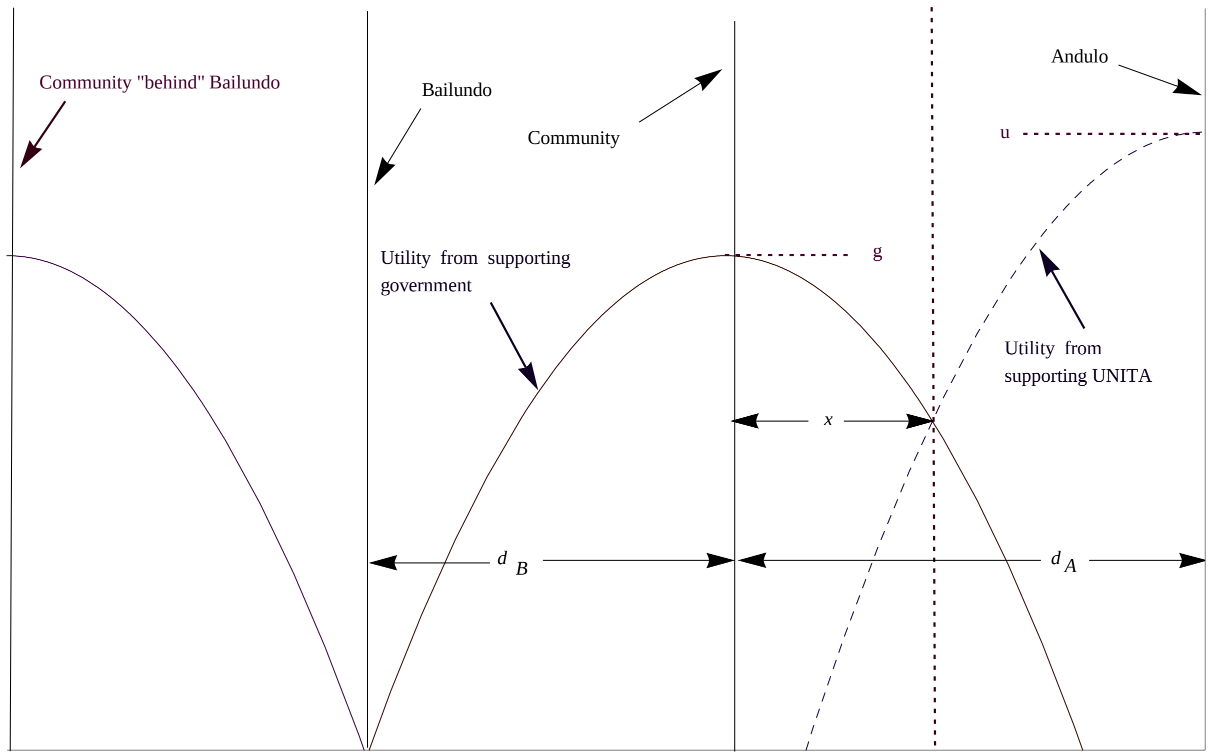


Figure 2.1: An illustration of the Hotelling spatial competition model underlying the first-stage reduced forms. If the government spends g in a given community, utility will be equal to $g - tx^2$ for a potential government supporter situated at distance x from that community; t parameterizes the degree of spatial differentiation.

Consider a community located somewhere between Andulo and Bailundo, at distance d_A from Andulo, and d_B from Bailundo. Assume that the government spends g in the community in the context of the FAS program. For simplicity, we assume quadratic "transportation costs", in which an individual located at distance x from the community that receives the government expenditures has utility given by:

$$U = g - tx^2,$$

where t parameterizes the degree of horizontal differentiation.

A graphical illustration of the model is provided in Figure 2.1, where the inverted

U-shaped utility functions reflect the quadratic transportion costs, which leads utility to be decreasing and concave as one moves away from the community being treated by the program. UNITA is located in Andulo. For areas lying between Andulo and the community treated by the FAS program, the dotted curve represents the utility of individuals who support UNITA. Given that the distance separating the community from Andulo is equal to d_A, an individual located at distance x from the community will be located at distance $d_A - x$ from Andulo.

An individual will be indifferent between supporting the government and UNITA when:

$$g - tx^2 = u - (d_A - x)^2,$$

where u is the level of utility of an individual located in Andulo proper and who supports UNITA. Solving the previous expression yields:

$$x = \frac{g - u + td_A^2}{2td_A}.$$

This is represented in Figure 2.1 by the intersection of the government's solid and UNITA's dotted utility curves. Individuals situated between the community and this intersection support the government, while those located between the intersection and Andulo support UNITA. When $g = u$, the population is evenly split between the government and UNITA. In Figure 2.1 we depict a situation in which $g < u$.

We assume that, for manifest security reasons, the government wishes to keep the political support of the population between the community in question and its bases in Bailundo. If we normalize the reservation level of utility of the consumer to zero (which corresponds to the horizontal axis in the Figure), such support between Bailundo and the community can be guaranteed by ensuring that:

$$g - td_B^2 \geqslant 0. \tag{2.1}$$

If the cost government expenditures is given by $\frac{1}{2}g^2$, where the convexity of the function reflects waste and the effect of rent-seeking activities, the government's optimization problem is then given by:[5]

$$\max_{\{g\}} \frac{g-u+td_A^2}{2td_A}+d_B-\frac{1}{2}g^2 \quad s.t. \quad td_B^2-g\leqslant 0,$$

which yields:

$$g^*=\begin{cases}\frac{1}{2td_A} & \text{for } \frac{1}{2t^2d_B^2}>d_A \\ td_B^2 & \text{otherwise}\end{cases}. \tag{2.2}$$

Note that for the government to maintain support amongst the population located between communities situated "behind" Bailundo and Bailundo itself, only the inequality given in (2.1) need be satisfied. In Figure 2.1 we depict a situation in which $\frac{1}{2td_A}=td_B^2$.

The implications of this highly stylized model of political competition are that FAS expenditures in a given community are: (i) *decreasing* in the distance d_A to Andulo in communities that are particularly close to this locality, (ii) *increasing* in the distance d_B to Bailundo for communities that are located far from Andulo. The model therefore suggests that the reduced form for treatment status, in which $Z_c=[d_A,d_B]$, should be given by:

$$D_c^*=X_c\pi_X+d_A\pi_A+d_B\pi_B-\lambda_c, \tag{2.3}$$

where $\pi_A<0$ and $\pi_B>0$. Whether our theoretical model provides the required identification, and whether it roughly corresponds to what is observed in the data can readily be tested by estimating (2.3) and examining both the significance and the signs of the coefficients associated with the distances separating each community from Andulo and Bailundo, respectively.

[5] Given that our goal is to characterize government spending behavior, we take u as being fixed. Allowing UNITA to choose u in an optimal fashion does not qualitatively modify the results of the model, since the expression for optimal government expenditures g^* given below in (2.2) is independent of UNITA behavior.

2.7 Empirical specification

Let i denote children, h households and c communities; N will denote sample size. The basic model that we are seeking to estimate is most easily understood to be given by:

$$Y_{ihc} = X_{ihc}\alpha + D_c\beta + \lambda_c\sigma + \varepsilon_{ihc}, \tag{2.4}$$

where Y_{ihc} is the $N \times 1$ vector associated with the outcome of interest (such as child health or household expenditures, in which case we shall write Y_{hc}), X_{ihc} is a $N \times K_X$ matrix of child, household and community control variables, D_c is an $N \times 1$ dummy that is equal to one when community c is treated by the FAS and zero otherwise, λ_c represents community level unobservables, and ε_{ihc} is a disturbance term. We assume that treatment status is given by a latent index model:

$$D_c^* = W_c\pi - \lambda_c = [X_c, Z_c]\,\pi - \lambda_c, \tag{2.5}$$

$$D_c = \mathbf{1}\left[D_c^* \geqslant 0\right], \tag{2.6}$$

where $W_c = [X_c, Z_c]$ is a $N \times K_W$ matrix of community-level covariates that determines treatment status, which we assume for the time being to be orthogonal with respect to λ_c and ε_{ihc}, $\mathbf{1}[.]$ is the indicator function that is equal to one when the inequality in square brackets is satisfied and zero otherwise, X_c is a matrix of covariates that determines treatment status and that also appear in the outcome equation, and Z_c is a matrix of variables that determine treatment status but that do not appear in (2.4). The "minus" sign in front of λ_c in (2.5) is a standard notational convention. In this "common factor" specification, the scalar parameter σ in (2.4) determines the "degree of endogeneity" of treatment status. Our purpose is to consistently estimate β, which measures the impact of treatment by the social program on our outcome variable.

In the context of the implementation of social programs such as the FAS, program initiators will usually favor communities perceived to be particularly poor. If all community-level covariates that determine treatment status are included in X_{ihc}, then we are in a situation of "selection on observables" and the system we are seeking to estimate reduces to:

$$\begin{aligned} Y_{ihc} &= X_{ihc}\alpha + D_c\beta + \varepsilon_{ihc}, & (2.7)\\ D_c^* &= X_c\pi - \lambda_c, & (2.8)\\ D_c &= \mathbf{1}\left[D_c^* \geqslant 0\right]. & (2.9)\end{aligned}$$

Estimation of (2.7) by OLS will then provide an unbiased estimate of β.

On the other hand, if observables included in X_{ihc} do not adequately control for those factors that affect treatment status, community-level unobservables λ_c that also determine treatment status will remain in (2.4), as parameterized by σ. If $\sigma > 0$, as is highly likely in the case of a poverty-reduction social program such as the FAS, the OLS estimate of β will be biased *downwards* because of the negative correlation between D_c and λ_c in (2.4) . If this bias is sufficiently large relative to the true value of β, it may lead to the least squares estimate of β being *negative*. In this case, the only solution is estimation by instrumental variables. More precisely, we must identify a matrix Z_c of excluded instruments that (i) are a significant determinant of treatment status in (2.5), and (ii) can plausibly be excluded from (2.4).

2.7.1 Data

The estimates presented in this chapter are based on the 1999-2000 *Inquérito aos agregados familiares sobre despesas e receitas* ("national household survey on expenditures and incomes", henceforth, IDR 2000), which was collected in the provinces of Cabinda, Luanda, Lunda Norte, Benguela, Namibe, Huila and Cunene. Given the unstable security situation at the time, the survey is roughly representative of areas of Angola under effective government control. The IDR 2000 includes information on household composition, expenditures, and child health.[6] It uses a stratified sampling design in which 12 households were surveyed in a random fashion in each **850** *aldeia* (village) in rural areas and *bairro* (neighborhood) in urban areas, in 50 *comunas.* Summary statistics for the two response variables and selected covariates for the 10,117 households and 8,328 children under the age of 5 included in the IDR 2000 data are presented in Table 2.1.[7] The survey was carried out by the *Gabinete de monitarização das condições de vida da população*, *Instituto nacional de estatística* (INE), *Ministério do planeamento* (MINPLAN). An interesting analysis of the situation in Luanda using the IDR data is provided by Rodrigues (2003); the only other uses of the survey to date, to the best of our knowledge, are provided by UN (2002) and Tvedten (2002).[8]

The IDR data are particularly appropriate for testing the impact of the FAS, given that they correspond almost perfectly to the provinces in which there are communities treated by the FAS I and II. As such, our estimates will be essentially based on within-province differences between treatment and control communities, and will, if there are spillovers from treated to untreated communities, provide a lower bound on the impact of the program. Only Kwanza Sul and Bengo were treated by the FAS I and II and were not included in

[6] We thank Lidia Galeano Germain for helping us understand the structure of the survey, as well as for filling in several gaps in the data. Formally speaking, the IDR data have not yet been officially released by the INE.

[7] Other data collected in Angola during this period includes the *Inquérito Prioritário sobre as Condições de Vida aos Domicílios* (IPCVD), which was conducted in 1995, and focused only on urban areas, in Benguela, Cabinda, Catumbela, Luanda, Lobito, Lubango and Luena. It is therefore sometimes known as the "urban poverty study." For our purposes, however, it was carried out too soon after the initial deployment of the FAS program. It also has relatively limited geographical coverage. Another potential source of data is constituted by the *Inquérito sobre a Disposição e Capacidade no Pagamento dos Serviços Sociais Básicos* (IDCPSSB) which was carried out in 1998 in Huambo, Huíla, Luanda and Uíge. While the main focus was the ability and willingness of households to pay for basic social services, it also collected data on household budgets. According to UN (2002), the IDCPSSB was probably the most representative survey on household expenditure conducted to date, with 55% of the population sample in urban areas (including 24% in Luanda) and 45% in rural areas. The survey found an average per capita expenditure of $2 a day. On these data, also see USAID (2002).

[8] A final source of microeconomic data for Angola during this period is provided by the 2001 *Multiple Indicator Cluster Survey*, (MICS 2001), carried out by UNICEF, which collected information on child anthropometrics, as well as on health and schooling issues, but no information on household budgets (a 1996 version of the survey also exists; see UNICEF (2003) for details on the most recent survey). Given the absence of household expenditure data, we have prefered to stick with the IDR. On child anthropometrics in Angola, apart from the MICS-based documents already cited, see GOA (2005). On educational attainment in Angola, see Bethke and Braunschweig (2003).

the IDR survey data, while only Lunda Norte is in the survey data without having any FAS-treated communities.

Our conflict intensity variable is based on the painstaking work of Ziemke (2007) who, based on archives, libraries and news agency files (a total of 186 sources from over twenty countries were involved), constructed a database of 9,216 individual battle and massacre events that took place in the Angolan war over a forty-one year period (1961-2002), as reported in the press. She provides the number of victims and the geographical coordinates for each event, allowing us, for each location included in the IDR 2000 dataset (for which we obtained the coordinates using Google Earth), to calculate the number of war-related casualties within a given radius of the community, in a given year. We focus on four different radii: 1, 5, 10 and 20 kilometres.

The final source of data is provided by the administrative records of the FAS, which allows us to identify which communities received funding from the FAS, and what form this funding took (education, health, water and sanitation, income generating activities and productive activities).

2.7.2 Treatment status

First-stage reduced forms of the determinants of treatment status corresponding to equation (2.3) are presented in the first two columns of Table 2.2. Column (1) presents linear probability estimates that correspond to the first-stage reduced form of the linear IV results for log household expenditures per adult equivalent that we present in section 2.8.1. Column (2) presents the corresponding first-stage reduced form for the child anthropometrics response variable. Column (3) presents the corresponding probit estimates that we will use in our analysis of the marginal treatment effect in section 2.8.4. We include the matrix of covariates X_{ihc} so as to avoid a "forbidden regression" problem in what follows. Covariates include child sex, age (in months) and age squared (for the child anthropometrics results), log household size, the dependency ratio, 8 dummies for the educational attainment of the household head, the household head's gender, 12 dummies for the type of water source used by the household (for the child anthropometric results), 12 dummies for the household head's language (which should control for ethnic-specific heterogeneity), six dummies for the household head's sector of employment, a rural-urban dummy, 7 provincial dummies, as well as the distance to Luanda and to the provincial capital. Standard errors are clustered at the *comuna* level in order to account for common shocks affecting all observations within a given *comuna* . As spelled out in our theoretical model in section 2.6.2, the distances of the community from Andulo and Bailundo are the two excluded instruments. Several aspects of the results are worth noting.

First, the likelihood of treatment by the FAS is a decreasing function of the distance of the community with respect to Andulo. Second, treatment by the FAS is an increasing function of the distance of the community with respect to Bailundo. Both of the point estimates are highly significant at the usual levels of confidence.[9] Third, the results are in agreement with the predictions of our spatial model of program deployment suggesting, at the very least, that treatment status is not neutral with respect to the location of these two communities.

While it is of course possible that the location of a community with respect to Andulo and Bailundo could be correlated with unobservables that affect our response variables, this is highly unlikely in that neither of these locations is either particularly important in economic terms, on the one hand, or isolated, on the other. In both cases the direction of the ensuing bias in the IV estimates is uncertain, since it depends not only upon the covariances between treatment status and our exclusion restrictions (which are known from

[9] Note that if we add the distance to Jamba (Cuando Cubango province) and the distance to Lumbala (Moxico province), two other locations of UNITA bases during this time period, as additional exclusion restrictions, the previously estimated coefficients remain statistically significant, with that associated with Bailundo doubling in size. The likelihood of treatment is increasing in the distance of the community from Jamba and decreasing in the distance from Lumbala. Note also that the linear IV results presented in section 2.8.1 remain both qualitatively and quantitatively unchanged when we use this extended instrument set. We prefer to confine ourselves to the more parsimonious instrument set since, in our opinion, our identification strategy is more transparent in this case. Results for the extended instrument set are, of course, available upon request.

the results presented in Table 2.2), but also upon the sign of the covariance between our response variables and the distances separating a community from Andulo and Bailundo. Nevertheless, in order to minimize the eventuality of such location-induced biases, we include, amongst the included covariates, measures of the distance separating the community from Luanda (the national capital) and the corresponding provincial capital.

2.8 Empirical results

2.8.1 Linear instrumental variables results

Table 2.3 presents instrumental variables estimates of the impact of treatment by the FAS on log household expenditures per adult equivalent and child height-for-age $z-$scores (HAZ). As shown by the coefficient reported in column (1), living in a FAS-treated community yields a 48.5% increase in household expenditures per adult equivalent, with an associated standard error of 0.175. The corresponding figure for total household income is 60.1% ($se = 0.232$). The coefficient reported in the uppermost portion of column (2) indicates that treatment by the FAS is associated with an increase of 0.359 in the HAZ ($se = 0.166$).

Both of these numbers are substantial, though the expenditures per adult equivalent result becomes understandable once the appallingly low initial level of income, caused in part by 25 years of civil war (and an additional 14 years of colonial-period struggles) is taken into account. The impact on child anthropometrics, which represents a reduction of 22.7% in the average height-for-age shortfall of the children included in the sample, is a relatively rare indication of the success of a social fund in terms of child health. On the other hand, the almost exclusive focus of existing research into social funds on educational outcomes (as summarized in section 2.3) may be in part to blame for this impression.

In the lower portion of the Table, we report the corresponding OLS-based coefficients. When no covariates are included, the coefficient reported in column (1) implies a *fall* of 47.4% in household expenditures per adult equivalent (which is significant at the 5%

confidence level), while when covariates are included, the coefficient increases to -0.055 ($se = 0.085$). This configuration of IV and OLS point estimates is what one would expect if the model sketched in equations (2.4), (2.5) and (2.6) holds: the OLS coefficient is biased downwards because of the negative correlation between treatment status D_c and community level unobservables λ_c that also affect the response variable; adding covariates X_{ihc} that are correlated with treatment status reduces this bias by purging part of the correlation between D_c and λ_c.[10] Finally, the putative "true" positive effect of treatment by the program is only uncovered once *exogenous* variation in treatment status is induced by appealing to excluded instruments. Much the same is true of the OLS estimates for the HAZ.

2.8.2 Observable sources of heterogeneity

From the policy perspective, heterogeneity in program impact is (or should be) a key aspect of any evaluation. For example, does the positive on HAZ that we attribute to treatment by the program stem from its effect on boys or from its effect on girls? Is the impact on children in poorer households, and in particular outside of Luanda or in rural areas, greater? Is program impact different for children who belong to households displaced by the war? Are there differential effects by age category (children in the 37 to 59 month range often suffer from greater nutritional shortfalls than younger children)?

We begin our exploration of impact heterogeneity by including interaction terms between the treatment status dummy and each of these categories. We follow Blundell, Dearden and Sianesi (2004) by instrumenting each of the multiplicative terms by our two exclusion restrictions multiplied by the appropriate covariate (the child's sex, a dummy indicating the child lives in Luanda, etc.), so as to avoid "forbidden regression" concerns (see, e.g. Wooldridge, 2002). As we can see in the Table 2.4, we find no statistically significant

[10]For the HAZ, both OLS coefficients are negative, but the results with covariates is statistically significant while without is not, at usual levels of confidence.

difference in the impact of treatment by the program for (i) boys versus girls, (ii) for children living in Luanda versus children living in other provinces, (iii) for children living in rural areas versus those living in urban areas, (iv) for children stemming from households that were displaced because of the war versus those that were not, (v) for younger children versus older children. This is in constrast to results reported by Levy and Ohls (2007) for a similar program in Jamaica where heterogeneity of impact based on sex, age and region of residence was found (though of course Jamaica hardly qualifies as a conflict setting). The policy implication, at least for the very specific Angolan context, is that there is no statistically-motivated reason to discriminate amongst these different categories of children in order to increase program effectiveness.

2.8.3 Possible transmission channels

Though we have identified a relatively homogeneous impact of treatment by the FAS I and II on HAZ $z-$scores, the question remains of the transmission channel through which this result obtains. In Table 2.5 we replace our HAZ $z-$score outcome variable with various *intermediate* outcomes that could lie behind our main result. Several findings are worth noting. First, the probability that the child is ill is not significantly affected by treatment, as is the likelihood that the child visited a healthcare center during the past three months: both of these variables are extremely rough measures of the use of healthcare facilities and it is perhaps not suprizing that no results of interest are found. Second, household expenditures on medical visits as well as household expenditures on medicines for children are significantly higher in FAS-treated communities: for expenditures on medicines for children, the difference in conditional means is equal to roughly 76%. Third, treatment by the FAS is associated with better access to clean water, whether this is measured by distance or by access times. Below 1,000m distances and below 60 minute access times are generally increased by exposure to treatment by the FAS,whereas above 1,000m distances and above 60 minute access times are generally reduced (the significance of the individual coefficients varies, though the pattern is relatively clear). Fourth, there appears to be no significant difference between treated and untreated households in terms of access to sanitation.

We do the same for intermediate outcomes corresponding to our observed increase in household expenditures per adult equivalent. In Table 2.6, three results are worth noting in this case. First, treatment by the FAS is associated with an 11% increase in the likelihood of the household head having working during the preceding week. Second, treatment by the FAS is associated with a mean increase of 1.33 hours in the number of hours worked by the household head during the preceding day, while treatment is associated with no increase in the number of days worked during the previous week or the number of months worked during the previous year. Third, treatment by the FAS is associated with a 57.6% increase in revenues stemming from the household head's principal activity and a corresponding increase of 19.2% in revenues from his secondary activity (though this result must be interpreted with caution given the much smaller sample size). The impression is that FAS treatment increases labor supply on a daily basis without increasing overall labor supply in terms of the days worked.

The upshot is that the effect of the FAS on HAZ $z-$scores appears to operate through improved access to clean water and on household medical expenditures, whereas the increase in household expenditures by adult equivalent appears to operate through a short-run increase in income accruing to the head's primary and secondary activities.

2.8.4 Unobservable heterogeneity and the marginal treatment effect of the FAS

An important question in the policy arena is how treatment by social programs such as the FAS is affected by the intensity of conflict. Does greater conflict intensity increase or reduce the benefits that accrue from treatment? The answer to this question is a key element in deciding whether social programs should be deployed in conflict situations, or whether they may actually be counterproductive, for example if they generate additional rents that the parties in conflict may fight over. Before examining this question, however, we must see whether the statistically significant treatment effects uncovered above using standard linear instrumental variables techniques are constant, or whether they vary with unobservables that determine treatment status. In order to do so, we appeal to the local instrumental variables (LIV) estimator of the marginal treatment effect pioneered by Heckman and Vytlacil (1999) and more fully described in Heckman and Vytlacil (2005) and Heckman, Urzua, and Vytlacil (2006)

The point of departure is a pair of potential outcome equations: Y_{1ihc} corresponds to the outcome when $D_c = 1$, whereas Y_{0ihc} corresponds to $D_c = 0$; all parameters and, most importantly, the disturbance terms, are allowed to differ by equation:

$$Y_{1ihc} = X_{ihc}\alpha_1 + \beta + \lambda_c\sigma_1 + \varepsilon_{1ihc}, \tag{2.10}$$

$$Y_{0ihc} = X_{ihc}\alpha_0 + \lambda_c\sigma_0 + \varepsilon_{0ihc}. \tag{2.11}$$

The system is completed by appending our latent index model given by (2.5) and (2.6). In particular, letting λ_c be distributed according to the standard normal density with *cdf* $\Phi(.)$, the propensity score is then defined as:

$$P(W_c) = \Pr\left(W_c\pi \geqslant \lambda_c\right) = \Phi\left(W_c\pi\right),$$

and treatment status is determined by:

$$D_c = \mathbf{1}\left[\Phi(\lambda_c) \leqslant P(W_c)\right]. \tag{2.12}$$

Note that, by the *Probability Integral Transform*, we can, without loss of generality, arbitrarily normalize the disturbance term λ_c so that it is distributed according to the uniform density over the unit interval, implying that treatment occurs when $W_c\pi \geqslant U_{Dc}$, where U_{Dc} ~Unif$[0,1]$, thereby preserving our notation for the propensity score as $P(W_c)$.

In section 2.7.2 (see Table 2.2, columns (1) and (2)), we presented first-stage reduced forms of the determinants of treatment status using the linear probability specification that corresponds to the application of standard instrumental variables techniques. In column 3 of Table 2.2, we estimate (2.12), which corresponds to a probit specification. Results are very similar to those reported earlier in columns (1) and (2). In particular, the distances to Andulo and Bailundo remain highly significant determinants of treatment by the FAS, and the signs predicted by our theoretical model remain. Results are qualitatively unchanged if we replace the probit specification by a logit. Histograms for the estimated propensity scores, for treated and untreated observations, are presented in Figure 2.3: the region of common support is given by the $[0.0296, 0.9917]$ interval.

In order to obtain an expression in terms of observables (Y_{1ihc} and Y_{0ihc} are never simultaneously observed for a given individual), note that the observable outcome Y_{ihc} can be expressed as:

$$Y_{ihc} = D_c Y_{1ihc} + (1 - D_c) Y_{0ihc}. \tag{2.13}$$

Substituting (2.10) and (2.11) into (2.13) then yields:

$$\begin{aligned} Y_{ihc} &= X_{ihc}\alpha_0 + D_c\beta + D_c X_{ihc} (\alpha_1 - \alpha_0) \\ &\quad + \lambda_c\sigma_0 + \varepsilon_{0ihc} + D_c \left[\lambda_c (\sigma_1 - \sigma_0) + (\varepsilon_{1ihc} - \varepsilon_{0ihc})\right]. \end{aligned} \tag{2.14}$$

It is worth noting that this equation collapses to the usual case (given above by equation (2.4)) that can be addressed using standard instrumental variables techniques (as with the results presented in section 2.8.1), when $\alpha_1 = \alpha_0 = \alpha$, $\sigma_1 = \sigma_0 = \sigma$ and $\varepsilon_{1ihc} = \varepsilon_{0ihc} = \varepsilon_{ihc}$. This observation is the basis for Heckman's test for the presence of essential heterogeneity, to which we now turn.

Taking the conditional expectation of (2.14) yields:

$$E\left[Y_{ihc}\right| X_{ihc}, P(W_c)] = X_{ihc}\alpha_0 + P(W_c)X_{ihc}\left(\alpha_1 - \alpha_0\right) + K(P(W_c)), \tag{2.15}$$

where:

$$K(P(W_c)) = P(W_c)\beta + \left[\lambda_c\left(\sigma_1 - \sigma_0\right) + \left(\varepsilon_{1ihc} - \varepsilon_{0ihc}\right)\right| \Phi(\lambda_c) \leqslant P(W_c)\right] P(W_c).$$

In the absence of essential heterogeneity, $P(W_c)X_{ihc}\left(\alpha_1 - \alpha_0\right)$ drops out of equation (2.15) and $K(P(W_c))$ reduces to $P(W_c)\beta$, leaving a specification that is *linear* in the propensity score.

That this is definitely not the case for log household income per adult equivalent is illustrated in Panel A of Table 2.8, which presents a semiparametric estimate, over the entire unit interval, of $K(P(W_c))$ in equation (2.15), which allows for a completely general relationship between the outcome and the propensity score.[11] A nonparametric test of the linearity of $K(P(W_c))$ in $P(W_c)$ rejects the null with an extremely low $p-$value, as reported in Table 2.3. The same is true when we replace the semiparametric specification of $K(P(W_c))$ with a fourth order polynomial in the propensity score, and when we add the interaction terms $(P(W_c)X_{ihc})$ between the covariates and the propensity score.

In contrast, estimates for child height-for-age $z-$scores fail to reject the absence of essential heterogeneity, as illustrated in Panel B of Table 2.8, which yields essentially a straight line for $K(P(W_c))$. To the extent that our identification strategy is valid, this suggests that the results presented for child anthropometrics in column (2) of Table 2.3 provide a reasonably faithful picture of the impact of treatment status on height-for-age $z-$scores.

The null of the absence of essential heterogeneity being strongly rejected for household

[11] We use a specification based on cubic thin plate splines with the smoothing parameter chosen automatically by restricted maximum likelihood (REML).

expenditures by our data implies that it is not clear what the linear instrumental variables results reported earlier actually represent. This is because, in the presence of essential heterogeneity, standard treatment parameters —such as the average treatment effect (ATE), the effect of treatment on the treated (TT) and the effect of treatment on the untreated (TUT)— differ. For child health, our results point towards all of these treatment parameters being equal. Estimating the marginal treatment effect (MTE) using LIV will allow us to distinguish between these different treatment parameters for household expenditures. It will also allow us to estimate, within the statistical structure given by this Roy model, how conflict intensity affects the impact of the FAS on household expenditures.

In order to do so, we begin by extracting the linear part of $E\left[Y_{ihc}\right| X_{ihc}, P(W_c)]$, using our estimate of (2.15), and purging it from Y_{ihc}. This yields:

$$\widetilde{Y}_{ihc} = Y_{ihc} - X_{ihc}\widehat{\alpha}_0 - P(W_c)X_{ihc}(\widehat{\alpha_1 - \alpha_0}).$$

We then obtain an estimate of $\frac{\partial \widetilde{Y}_{ihc}}{\partial p_c}$ by regressing $\widetilde{Y}_{ihc}$ nonparametrically on $P(W_c)$ over the region of common support. The MTE is then given by:

$$\begin{aligned} \Delta^{MTE}\left(x_{ihc}, u_{Dc}\right) &= \left.\frac{\partial E\left[Y_{ihc}\right| X_{ihc} = x_{ihc}, P(W_c) = p_c]}{\partial p_c}\right|_{p_c = u_{Dc}}, \\ &= x_{ihc}(\widehat{\alpha_1 - \alpha_0}) + \left.\frac{\partial K(P(W_c) = p_c)}{\partial p_c}\right|_{p_c = u_{Dc}}, \\ &= x_{ihc}(\widehat{\alpha_1 - \alpha_0}) + \left.\frac{\partial \widetilde{Y}_{ihc}(P(W_c) = p_c)}{\partial p_c}\right|_{p_c = u_{Dc}}. \end{aligned}$$

The nonlinear portion ($\frac{\partial K}{\partial p_c} = \frac{\partial \widetilde{Y}_{ihc}}{\partial p_c}$) of the MTE for log household income per adult equivalent is represented by the solid curve in Figure 2.4 (the dotted curves correspond to the ± 2 standard error confidence bands).[12] The solid horizontal line in the figure corresponds to *minus* the linear portion of the MTE ($x_{ihc}(\widehat{\alpha_1 - \alpha_0})$), evaluated at the mean sample

[12]Note that derivative estimation is an ongoing area of research in nonparametric statistics. A good primer on the topic is provided by Newell and Einbeck (2007).

values of x_{ihc}. In terms of understanding the overall magnitude of the MTE, this horizontal line will therefore shift up or down depending upon the subset of observations over which $x_{ihc}(\widehat{\alpha_1 - \alpha_0})$ is evaluated. An intuitive manner of interpreting the figure is therefore that different subsets of observations correspond to different vertical intercepts of this horizontal line, with each intercept corresponding to a particular normalization.

As is evident upon inspecting Figure 2.4, the MTE displays an inverted$-U$ pattern. For both low and high values of u_{Dc}, the MTE is negative (though not statistically significant for low values of u_{Dc}), whereas it becomes positive for intermediate values of u_{Dc}. For households living in communities with unobservables such that it was highly likely that they received treatment by the FAS, which corresponds to low values of u_{Dc} (which in turn corresponds to low values of λ_c in the latent index model given by (2.5)), the MTE is negative and insignificant. For communities with unobservables such that it was highly *unlikely* that they received treatment by the FAS, the MTE is both negative and statistically significant at the usual levels of confidence. Figure 2.4 makes it clear that the substantial treatment effect estimated using linear IV techniques is attributable in large part to communities with u_{Dc} lying roughly in the $[0.15, 0.75]$ interval. Figure 2.4 also makes clear that while no formal targeting mechanism was set up in the context of program deployment, there was a very high probability of the FAS bypassing communities whose benefits would have been negative. On the other hand, unobservables making it highly likely that a community would be treated were associated with MTEs that were effectively zero.[13]

[13] Using the weights provided in Heckman (2001), we can derive the three standard treatment parameters by computing the appropriate weighted averages of our estimates of the MTE, though these results are not particularly informative, given the highly nonlinear aspect of the curve. Note that these are empirical versions of the treatment parameters, in that the region of common support is not the full unit interval. Given that there is a large portion of the empirical MTE that is negative when evaluated at the mean sample value of conflict intensity, the estimated ATE is negative ($\text{ATE} = -0.067, se = 0.197$) though it is not statistically significant at the usual levels of confidence. The same is true for the TT ($= -0.180, se = 0.212$) and TUT ($= -0.147, se = 0.170$).

2.9 Conflict intensity and the marginal treatment effect

It is important to keep in mind that our purpose here is *not* to consistently estimate the impact of conflict intensity on our response variables.[14] Instead, our purpose is to assess whether the impact of the FAS on our response variables is a function of conflict intensity. In the context of our estimation strategy, assessing the effect of conflict intensity on the MTE is extremely straightforward. To see why, reconsider equation (2.15), in which we explicitly include conflict intensity, denoted by C_c, as a covariate:

$$\begin{aligned} E\left[Y_{ihc}\right| X_{ihc}, P(W_c)] &= X_{ihc}\alpha_0 + C_c\delta_0 + P(W_c)X_{ihc}\left(\alpha_1 - \alpha_0\right) \\ &\quad + P(W_c)C_c\left(\delta_1 - \delta_0\right) + K(P(W_c)). \end{aligned} \tag{2.16}$$

The marginal treatment effect is then given by:[15]

$$\Delta^{MTE}\left(x_{ihc}, u_{Dc}\right) = c_c(\widehat{\delta_1 - \delta_0}) + x_{ihc}(\widehat{\alpha_1 - \alpha_0}) + \left.\frac{\partial \widetilde{Y}_{ihc}(P(W_c) = p_c)}{\partial p_c}\right|_{p_c = u_{Dc}}.$$

The impact of conflict intensity on the MTE is therefore given by $c_c(\widehat{\delta_1 - \delta_0})$, which we can directly measure through a semiparametric estimate of (2.16).[16] An important point, first noted by Heckman and Vytlacil (1999), and that is in contrast to standard econometric approaches, is that X_{ihc} may be correlated with the disturbances in the potential outcome equations: in other words, X_{ihc} need not be "exogenous" in any conventional definition of that term.[17]

[14] This is done in a companion piece.

[15] Of course, $\widetilde{Y}_{ihc} = Y_{ihc} - X_{ihc}\widehat{\alpha}_0 - C_c\widehat{\delta}_0 - P(W_c)X_{ihc}(\widehat{\alpha_1 - \alpha_0}) - P(W_c)C_c(\widehat{\delta_1 - \delta_0})$ in this case.

[16] See Carneiro, Heckman, and Vytlacil (2003) who show, in exactly the same statistical setup, that individuals with good results on the Armed Forces Qualifying Test (AFQT) have their MTE from completing college (the response variable is log wage) shifted upwards.

[17] The key orthogonality assumption is that the three disturbance terms must be orthogonal with respect to the excluded instruments, *conditional* on the included covariates: $(\lambda_c\sigma_1 + \varepsilon_{1ihc}, \lambda_c\sigma_0 + \varepsilon_{0ihc}, \lambda_c) \perp Z_c \mid [X_{ihc}, X_{hc}, X_c]$.

Our conflict intensity variable is based on the painstaking work of Ziemke (2007) who, based on archives, libraries and news agency files (a total of 186 sources from over twenty countries were involved), constructed a database of 9,216 individual battle and massacre events that took place in the Angolan war over a forty-one year period (1961-2002), as reported in the press.[18] She provides the number of victims and the geographical coordinates for each event, allowing us, for each location included in the IDR 2000 dataset, to calculate the number of war-related casualties within a given radius of the community, in a given year. We considered four different radii: 1, 5, 10 and 20 kilometres, and present results that correspond to the 20km radius; results for the three other radii where qualitatively similar, with minor quantitative differences. Summary statistics for the conflict intensity variables are presented in the lower portion of Table 2.1.

In quantitative terms, the point estimate of $\widehat{\delta_1 - \delta_0}$ $(0.182, se = 0.026)$, reported in the lowermost portion of Table 2.7, implies that each additional $1,000$ casualties within 20 km of the community during the 1995-2000 period is associated with an 18.2% increase in the MTE of the FAS. These results indicate that while the impact of treatment by the FAS on household expenditures per adult equivalent was characterized by a great degree of heterogeneity driven by unobservables that determined treatment status, communities that bore the brunt of the fighting were the biggest beneficiaries.[19] In the Angolan context, at least, greater conflict intensity was associated with an increase in the benefits that local populations derived from treatment by this particular social fund.

[18]This work is part of the Armed Conflict Location and Events Dataset (ACLED).

[19]Note that we can find no specification in which the marginal impact of treatment status on HAZ is a statistically significant function of conflict intensity. For example, if we revert to standard IV techniques, allow for an interaction term between treatment status and conflict intensity, assume that both conflict intensity and the interaction term are jointly endogenous (and use the product of the distance variables as the required third exclusion restriction), the interaction term remains statistically indistinguishable from zero.

2.10 Concluding remarks

This chapter has considered the impact of a typical social fund –the FAS I and II– on child anthropometrics and household expenditures per adult equivalent. The specificity of our results stems from the fact that the social fund in question was operating during the final period of the Angolan civil war, and that the identification strategy used to identify this effect is motivated by the political geography of the conflict in question. Linear instrumental variables estimates, based on household survey data collected in 2000 indicate that le FAS a un impact positif et significatif sur height-for-age and les dépenses par équivalent adulte.

A somewhat suprizing aspect of our findings is that the impact of the FAS is found to be relatively homogeneous, and this whether we allow for differences based on the sex, residence, age, or displaced status of the children involved. In terms of program targeting, our results therefore suggest that there is no reason to priviledge particular classes of children for program treatment. Our analysis of the channels through which the impact of the program obtains indicate that it is increased household expenditures and better access to clean water that constitute the key.

In retrospect, the inverted$-U$ shaped MTE for household expenditures per adult equivalent is perhaps not so surprizing. Assume for argument's sake that program initiators would have been able to identify communities with particularly large MTEs. In this case, and if the purpose of targeting had been to concentrate deployment on communities which had the most to gain from the program while avoiding communities with particularly low (or negative) MTEs, one would expect to see a *downward-sloping* MTE curve: communities with unobservables such that they would have been highly likely (unlikely) to be treated would have large (small) MTEs.

That this is *not* the case indicates two things. First, the program was not targeted so as to maximize economic gains. Second, given that maximizing economic gains was not the goal, something else was the driving force behind program deployment, and our hypothesis in this chapter, crystallized in our identification strategy based on the political geography of the Angolan conflict, has been that it was political influence on the ground, particularly in areas that were on the frontline at the time of program initiation.

In policy terms, and while one must be cautious in generalizing from results based on a specific conflict situation, especially one as complex as Angola, our findings suggest at least three lessons. First, if our identification strategy is valid, and if, indeed, the significant positive effect of treatment by the program on child anthropometrics is neutral with respect to conflict intensity, there is no reason *not* to deploy social programs in areas of conflict when maintaining child health is the paramount goal.

Second, given that the marginal treatment effect of the program on household expenditures per adult equivalent is increasing in conflict intensity, there are good economic reasons for concentrating program deployment in areas where violence is (or has been) particularly severe, though this prescription must necessarily be tempered by the dangers to program staff that stem from such a form of geographical targeting.

Third, while it would be preferable on efficiency grounds to target the program so as to guarantee a downward-sloping MTE curve, thereby ensuring that it is highly likely that those communities with particularly large MTEs receive treatment, such a prescription is undoubtedly naive from the political standpoint: in situations of conflict, social programs necessarily become a tool associated with grassroots support in the field, and ignoring government political goals in the interests of efficiency may, ultimately, be counterproductive in terms of the welfare of the populations concerned.

Figure 2.2: The FAS I and II intervention areas, in purple. Orange provinces (Bie and Zaire) correspond to provinces included in the FAS III. Green provinces are currently being treated.

	mean	median	std.
Response variables: y_{ihc} or y_{hc}			
Child height-for-age (HAZ) $z-$score	-1.577	-1.620	1.433
Household expenditures per adult equivalent (Kwanzas)	390	249	477
Selected covariates: x_{ihc}, x_{hc} and x_c			
Child age (months)	28.09	28.00	17.10
Child is female	0.490	0.000	0.499
Household is Kimbundo	0.469	0.000	0.499
Years education of household head	4.537	5.000	2.056
Acces to water in the house	0.120	0.000	0.325
Household size	7.27	7.00	3.17
Distance to:			
Luanda, in 10 km	38	38	37
Provincial capital, in 10 km	8	3	13
Exclusion restrictions: z_c			
Distance to UNITA headquarters in Andulo, in 10 km	46	45	19
Distance to Bailundo, in 10 km	50	46	17
Conflict intensity: C_c			
Number of casualties, 1995-2000 (in thousands), within a radius of:			
1 km	0.016	0.000	0.044
5 km	0.020	0.000	0.048
10 km	0.038	0.000	0.059
20 km	0.102	0.025	0.140

Table 2.1: Descriptive statistics: response variables and covariates (8,328 observations for child-specific variables, 10,017 observations for household-specific variables).

Dependent variable:	Treatment status: D_c Equals one if the community was treated by the FAS I or II programs during the 1994-2000 period, and zero otherwise		
	Linear probability model		Probit
	Log household expenditures per adult equivalent response variable	Child height-for-age $z-$score (HAZ) response variable	Log household expenditures per adult equivalent response variable
	(1)	(2)	(3)
Exclusion restrictions Z_c			
Distance of community from (in 10 km):			
Andulo (Bie)	-0.143 (0.057)	-0.151 (0.077)	-0.814 (0.257)
Bailundo (Huambo)	0.100 (0.029)	0.071 (0.034)	0.560 (0.190)
Covariates $[X_{hc}, X_c]$			
Distance from Luanda	0.066 (0.040)	0.073 (0.054)	0.033 (0.014)
Distance from provincial capital	-0.017 (0.014)	-0.056 (0.009)	-0.005 (0.013)
Child variables included	no	yes	no
Household variables included	yes	yes	yes
Provincial dummies included	yes	yes	yes
$\overline{R}^2$ or pseudo-$\overline{R}^2$	0.671	0.680	0.750
$F-$statistic	F^{98}_{9902} [$p-$value] $= 209$ [0.000]	F^{75}_{8252} [$p-$value] $= 244$ [0.000]	

Table 2.2: The determinants of treatment status; 10,000 observations in columns (1) and (3), 8,327 observations in column (2). Standard errors clustered at the community level in parentheses.

Dependent variable Y_{hc} or Y_{ihc}	Log household expenditures per adult equivalent (1)	Child height-for-age $z-$score (HAZ) (2)
Instrumental variables estimates		
Treated by the FAS: D_c	0.485 (0.175)	0.359 (0.166)
H_0: D_c exogenous ($p-$value)	0.001	0.000
H_0: absence of essential heterogeneity ($p-$value)		
4th order polynomial	0.000	0.738
4th order polynomial + interaction terms	0.000	0.814
nonparametric	0.000	0.068
Excluded IVs. Distance from: Andulo, Bailundo	yes	yes
OLS estimates: covariates not included		
Treated by the FAS: D_c	-0.474 (0.237)	-0.128 (0.069)
OLS estimates: covariate included		
Treated by the FAS: D_c	-0.055 (0.085)	-0.185 (0.082)

Table 2.3: Instrumental variables estimates of the impact of treatment by the FAS I or II programs on log household expenditures per adult equivalent and child height-for-age z-scores; 10,000 observations in column (1), 8,327 observations in column (2); child- and household-specific covariates, distance to Luanda and to provincial capital, and province dummies included as appropriate. Standard errors clustered at the community level in parentheses.

Heterogeneous Treatment Effects on HAZ	
Boys	0.023 (0.123)
Girls	0.347 (0.394)
Luanda	0.477 (0.793)
Province other than Luanda	0.441 (0.649)
Urban Area	0.662 (0.632)
Rural Area	0.704 (0.514)
Child living in a displaced household	0.418 (0.408)
Child not living in a displaced household	0.508 (0.436)
Child is between 0 and 36 months old	0.282 (0.264)
Child is between 37 and 59 months old	0.535 (0.386)

Table 2.4: Heterogeneous of the impact of treatment by FAS program on child height-for-age z-scores (HAZ) along different demographic and location characteristics; 8,328 observations; child- and household-specific covariates, distance to Luanda and to provincial capital, and province dummies included as appropriate. The coefficient is the interaction between the treatment indicator variable with one demographic or location variable. Standard errors clustered at the community level in parentheses.

Health of children and access to medical care	
Child was ill in the previous week	0.005 (0.005) [8321]
Child has made the medical analysis in the last three months	0.161 (0.172) [8298]
Log expenditures for the consultation in hospital in previous month	1.381 (0.621) [6034]
Log expenditures for the drug of children in previous month	0.764 (0.382) [5847]

Access to water and improved Sanitation	
Distance to clean water:	
distance < 100 meter	0.159 (0.127) [7338]
100m < distance < 500m	0.076 (0.042) [7338]
500m < distance < 1000 m	0.030 (0.082) [7338]
1000m < distance	−0.113 (0.071) [7338]
Time to clean water:	
time < 15 minutes	0.162 (0.129) [7335]
15 minutes < time < 30 minutes	0.102 (0.043) [7335]
30 minutes < time < 60 minutes	0.044 (0.062) [7335]
60 minutes < time	−0.043 (0.031) [7335]
Household have improved sanitation	−0.012 (0.133) [8325]

Table 2.5: The impact of FAS program on intermediate outcomes in assessing the transmission channels of FAS on child height-for-age z-scores (HAZ); observation number in bracket; child- and household-specific covariates, distance to Luanda and to provincial capital, and province dummies included as appropriate. Standard errors clustered at the community level in parentheses

Impact of FAS on access to employment and household income	
Head of household having worked in the previous week	0.113 (0.085) [9021]
The number of working hours in previous working day	1.334 (0.458) [9021]
The number of working day in previous month	0.044 (0.095) [9021]
The number of month worked in the previous year	0.157 (0.281) [9021]
Log of income from main activity of head of household	0.576 (0.241) [9976]
Log of income from second activity of head of household	0.192 (0.027) [987]

Table 2.6: The impact of FAS program on intermediate outcomes in assessing the transmission channels of FAS on log household expenditures per adult equivalent; observation number in bracket; household-specific covariates, distance to Luanda and to provincial capital, and province dummies included as appropriate. Standard errors clustered at the community level in parentheses

Dependent variable Y_{hc}	Log household expenditures per adult equivalent
Selected covariates X_{hc} : $\widehat{\alpha}_0$	
Urban household	0.834 (0.043)
Female head	−0.071 (0.018)
Household size	−0.517 (0.013)
Dependency ratio	−0.039 (0.008)
Distance from Luanda	0.002 (0.000)
Distance from provincial capital	0.002 (0.000)
Conflict intensity C_c : $\widehat{\delta}_0$	0.002 (0.012)
Covariates interacted with propensity score $P(W_c)X_{hc}$: $\widehat{\alpha_1 - \alpha_0}$	
Distance from provincial capital	0.002 (0.000)
Conflict intensity $P(W_c)C_c$: $\widehat{\delta_1 - \delta_0}$	0.182 (0.026)

Table 2.7: Semiparametric estimate of the conditional expectation of household expenditures per adult equivalent: linear portion; 10,000 observations. Standard errors in parentheses.

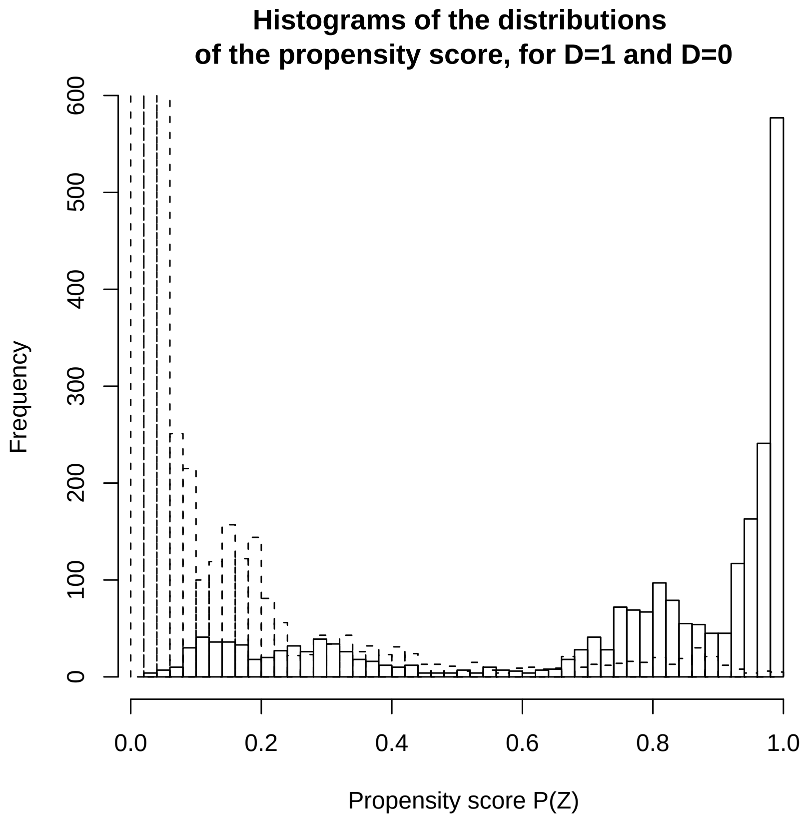
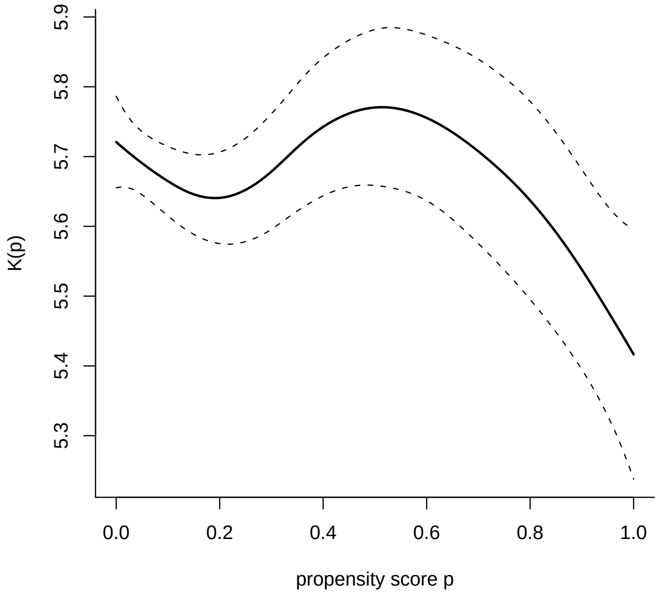
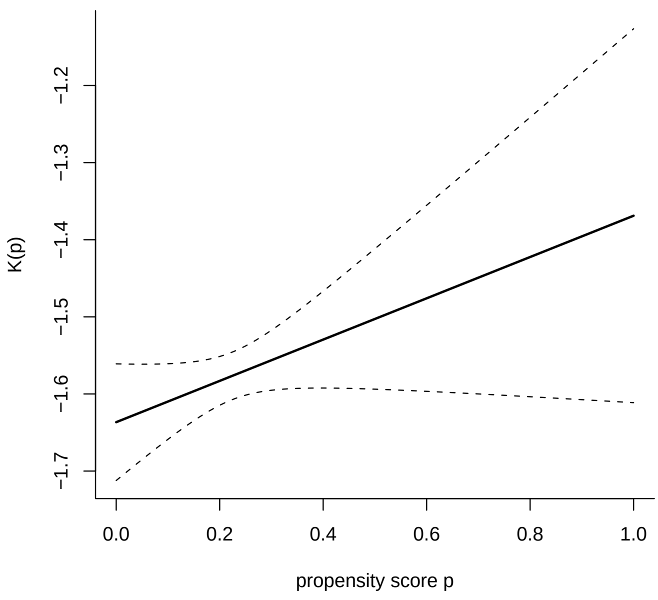


Figure 2.3: Histograms of estimated propensity score, for $D = 1$ (solid line) and $D = 0$ (dotted line). The region of common support is the $[0.0296, 0.9917]$ interval.



Panel A: Log household expenditures per adult equivalent.



Panel B: Child height-for-age $z-$score.

Table 2.8: The potential nonlinearity of K(p) as a function of the propensity score, for log household expenditures per adult equivalent (Panel A, 10,000 observations), and child height-for-age z-scores (Panel B, 8,327 observations) (dotted lines represent plus/minus 2 standard error confidence bands).

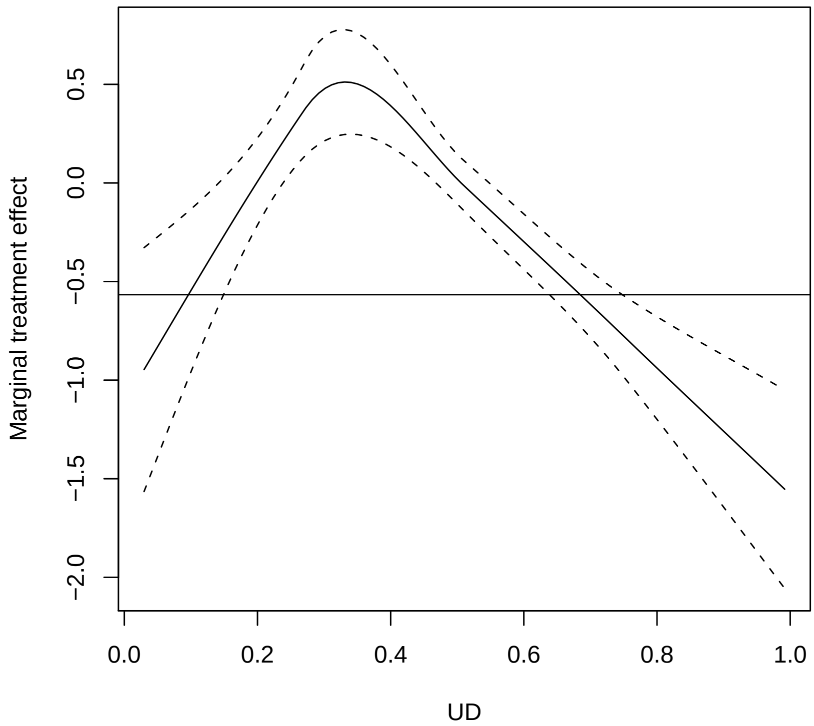


Figure 2.4: The marginal treatment effect of the FAS I and II on log household expenditures per adult equivalent (with ± 2 standard error confidence bands). Solid horizontal line corresponds to $-\overline{c}_c(\widehat{\delta_1 - \delta_0}) - \overline{x}_{ihc}(\widehat{\alpha_1 - \alpha_0})$, the linear portion of the MTE evaluated at its sample mean.

Chapter 3

Teacher Training and HIV/AIDS prevention in West Africa: Regression discontinuity design evidence from the Cameroon

[1]

[1] A version of this chapter was published under reference: "Teacher Training and HIV/AIDS prevention in West Africa: Regression discontinuity design evidence from the Cameroon", *Health Economics*, 19(S1), 36-54. September 2010

3.1 Introduction

Sub-Saharan Africa accounts for 70% of the annual number of new HIV infections (UNAIDS 2008). According to the Joint United Nations Programme on HIV/AIDS, HIV infections are spreading most quickly within youth populations. Young people aged 15–24 account for 45% of all new HIV infections in adults, and many young people still lack accurate, complete information on how to avoid exposure to the virus. The HIV epidemic cannot be reversed without strong, sustained success in preventing new HIV infections. As such, HIV is a major public health problem for many Sub-Saharan African countries, with Cameroon being the most affected in West Africa, a region that is typically less affected by the epidemic compared with East or Southern Africa.

Since HIV infections cannot be cured, effective prevention campaigns are currently the most important tools to control the epidemic. As a result, a substantial amount of financial resources is devoted to HIV/AIDS prevention programs. This is particularly true of school-based programs, since a large fraction of the population, in most African countries, attends either primary or secondary school. Most countries in sub-Saharan Africa are integrating HIV education into their secondary school curricula (Ramos and Siamatowe 2006). However, as documented in the surveys by Gallanta and Maticka-Tyndale (2004) and Kirby, Laris, and Rolleri (2007), rigorous evaluations of the effectiveness of such interventions are rarely undertaken, which is suprizing given the debate as to whether such programs can be effective in limiting the spread of HIV/AIDS amongst African youth.

Though there are by now many papers which evaluate school-based programs, almost all suffer from methodological limitations, including (i) the lack of a clear identification strategy, (ii) the reliance on self-reported sexual behavior instead of biological outcomes, (iii) low statistical power because of inadequate sample size, or (iv) inappropriate timing in terms of when the intervention being evaluated was carried out and when data on students were collected.

As noted by Gallanta and Maticka-Tyndale (2004), papers in which an identification strategy is clearly spelled out are relatively rare, though a number of randomized control trials (RCTs) stand out. Stanton, Li, Kahihuata, Fitzgerald, Neumbo, Kanduuombe, Ricardo, Galbraith, Terreri, Guevara, Shipena, Strijdom, Clemens, and Zimba (2009) is a typical example of an RCT applied to a school-based intervention. The authors find that sexually inexperienced Namibian youths aged 15 to 18 participating in an school-based AIDS education program reduce a number of self-reported risky behaviors, including sex with multiple partners during the last month and sex without a condom. However, no biological outcomes were considered. On the other hand, when biological outcomes, such as other STDs or teenage pregnancy are considered in RCTs, sample size is often so small that lack of statistical power is the issue. A typical example is Hayes, Changaluchab, Ross, Gavyole, Todd, Obasi, Plummer, Wight, Mabey, and Grosskurth (2005) who considered 10 treatment and 10 control communities in rural Tanzania.

The time which elapses between treatment and measurement of the potential effects is also often a problem. Kirby, Laris, and Rolleri (2007) review 83 studies (19 of which were undertaken in developing countries) that measure the impact of curriculum-based sex and HIV education programs on sexual behavior and mediating factors among youth under 25 years of age. Two thirds of the programs significantly improved one or more sexual behaviors, though the identification strategies are almost never clearly defined. Most of these programs involved some form of training, be it of teachers, peers or medical personnel. 41% of the studies involved surveys that were carried out within less than 3 months of treatment by the program, while 37% involved a 4 to 6 month time-span. Statistical power is also a problem, with an average of 2 treated and 3 control schools in the 83 studies considered, and samples of between 200 and 1,000 students. When panel data are used, simple (pre/post) difference results (and not difference-in-differences) are the norm.

Notable exceptions to this somewhat bleak picture in terms of the evaluation of school-based programs are provided by Duflo, Dupas, Kremer, and Sinei (2006) and Dupas (2009).

The former conduct a randomized evaluation of the impact of the three different school-based interventions, including teacher training, on a large sample of Kenyan schools and students. Their primary measure of the effectiveness of the interventions they consider is teenage childbearing, which is associated with unprotected sex. They also collected self-reported measures of knowledge, attitudes and behavior regarding HIV/AIDS. After two years, girls in schools where teachers had been trained were more likely to be married in the event of a pregnancy. However, the program had little impact on students' knowledge, attitudes, and behavior. It also had no statistically significant impact on the incidence of teen childbearing.

Based on a subsample of the Kenyan schools considered in Duflo, Dupas, Kremer, and Sinei (2006), Dupas (2009) uses a randomized experiment to test whether teenage girls with full information on HIV prevalence disaggregated by gender and age groups can change their sexual behavior. She finds that information on the relative risk of HIV infection by partner's age led to 1.5% reduction in the incidence of having started childbearing amongst treated girls compared to untreated girls (a 28% decrease). Self-reported sexual behavior data suggests that the mechanism underlying this result is substitution away from older (riskier) partners and towards protected sex with same-age partners.[2]

In this chapter we present quasi-experimental evidence of the impact on teenage childbearing and self-reported knowledge, attitudes and behavior of a school-based teacher training program in the Cameroon: the *EVF/EMP/VIH et SIDA*.[3] Our regression discontinuity

[2] There is also a related, growing, literature which uses experimental methods to examine other interventions that may modify behavior in ways that will contribute to reducing the spread of HIV/AIDS. Thornton (2008) constructs an experiment in which individuals in rural Malawi were randomly assigned monetary incentives to learn their HIV results after being tested. Distance to the HIV results centers was also randomly assigned. Without any incentive, 34% percent of the participants learned their HIV results. However, even the smallest incentive doubled that share. Sexually active HIV-positives who learned their results were three times more likely to purchase condoms two months later than HIV-positives who did not learn their results. Rosenblum, Jewell, Van Der Laan, Shiboski, Van Der Straten, and Padian (2009) investigated the effect of diaphragm and lubricant gel use in reducing HIV among women in South Africa and Zimbabwe. Though there was a much lower reported use of condoms in the intervention arm than in the control arm, the authors found a statistically significant protective effect of diaphragm/gell use. Finally, Frolich and Vazquez-Alvarez (2009) use DHS data and IV techniques to study how knowledge concerning one's serological status affects general knowledge concerning HIV.

[3] This rather cumbersome acronym corresponds to *Education à la vie familiale, en matière de population*

design (RDD) identification strategy exploits the natural experiment generated by the manner in which the program was deployed: teacher training on HIV/AIDS was introduced in secondary schools in towns with less than 4 secondary schools, whereas it was not in towns with 5 or more secondary schools. As noted by Lee and Lemieux (2010) in their recent survey article on RDD, this provides us with a "local randomized experiment." Our results indicate that the incidence of childbearing for 15-17 year old girls is between 7 and 10 percentage points lower for girls attending teacher training schools than those attending control schools. They are also significantly more likely to have used a condom during their last sexual intercourse. For 12-13 year old girls, the likelihood of self-reported abstinence and condom use is also significantly higher in treated schools, while the likelihood of having multiple partners is significantly lower.

The rest of this chapter is organized as follows. Section 3.2 provides background information on the Cameroonian context, including HIV prevalence, awareness, and government policy. Section 3.3 describes the program. Section 3.4 discusses the data and the estimation strategy. Section 3.5 presents the empirical results. Section 3.6 concludes, and presents some thoughts on the implications of our results for the role that can be played by teacher training as a means of fighting HIV/AIDS in the Cameroon.

3.2 Background on HIV/AIDS in the Cameroon

3.2.1 HIV prevalence

Data collected through the 2004 DHS-III show that 5.5% of all Cameroonian adults aged 15-49 years are HIV-positive. The rate for women in this age bracket is higher (6.8%) than for men (4.1%). Prevalence rates increase with age, from 1.4% for the 15-19 age bracket, to 5.5% for males and 7.9% for females in the 20-24 year age bracket. Prevalence rates are higher in urban areas, with a national average 6.7%, than in rural areas, where prevalence is

et au VIH/SIDA.

4.0%. Large regional disparities are also apparent. In the Northwest province, prevalence stands at 8.7%, in the Eastern province it is 8.6%, whereas in the North and Extreme North the prevalence rates are 1.7% and 2.0%, respectively.

Most observers in the Cameroon agree that the epidemic is partly driven by lack of knowledge concerning HIV and risky sexual behavior, including multiple partners, as well as occasional and especially paid sex. Such behavior is, according to the 2008 Cameroun Progress Report, particularly frequent in the 15-24 year age category. The MICS 2006 data, which include a focus on female reproductive health, show that only 32% of women aged 15 to 24 report knowing of at least two methods of preventing the transmission of HIV.

The 2004 DHS-III also shows that 55% of men report having used a condom during their last sexual encounter. This percentage is identical when one considers only the 15-19 year age bracket. For women, the corresponding figure is 41%, with slightly less risky behavior by younger women (15-19 years of age) for whom the figure is 47%.

Knowing one's serological status is also an important factor in containing the spread of HIV in that it allows one to protect one's partner. According to the 2004 DHS-III, only 10% of women had ever carried out an HIV test and knew their results. The corresponding figure is slightly higher for men and stands at 14%. If one restricts one's attention to the previous 12 months, the corresponding figures are 5% for women and 7% for men. Whether it be for females or for males, the lowest level of knowledge of one's HIV status is to be found in the youngest age category: 5% for females, and 3% for males. This suggests that there is significant room for improvement in terms of increasing voluntary counciling and testing (VCT) amongst lower age categories: school-based programs such as the one considered in this chapter could be particularly useful in this respect.

3.2.2 HIV/AIDS awareness

When the first case of HIV/AIDS was identified in the Cameroon in 1985, the initial response of the authorities was swift, with the almost immediate creation of the first *comité de lutte contre le SIDA* within the Ministry of public health; this was followed, on the basis of WHO guidelines, with the creation in 1987 of the *Programme national de lutte contre le SIDA* (PNLS). Two medium-terms plans were put into place during the 1988-1998 period. Because of poorly crafted social communication and a lack of coordination amongst various actors, they are widely held to have been failures according to Eboko (2006). According to data collected by the Ministry of public health in prenatal clinics, prevalence rates increased from 0.5% in 1987 to 10.8% in 2000.

Starting in 2000, attempts were made to remedy the situation, including the adoption of strategic and multisectoral plans during the 2000-2005 period with the financial support of the World Bank. The emphasis during this period was put on decentralizing the response to the epidemic. This has entailed the involvement of all sectors of Cameroonian society in the fight against HIV/AIDS. In administrative terms, decentralization has been crystallized, since 2006, in the creation of health districts that encompass 5,000 to 10,000 inhabitants, and which constitute the basic operational unit within which the fight against HIV/AIDS is carried out.

3.2.3 Government policy on HIV/AIDS education

Almost one in four Cameroonians was enrolled in the schooling system in 2004, for a total of 4 million individuals. According to 2004 *Annuaire statistique*, this is projected to increase to 5.5 million by 2015. As such, the Cameroonian educational system is one of the most important entry points for reaching the country's youth in terms of HIV/AIDS. Before the introduction of the teacher training program considered here, a variety of interventions had been implemented with the aim of reducing the prevalence of HIV in Cameroonian schools.

These included organizing school days devoted to STDs, HIV and AIDS, VCT services for students, the UNICEF-sponsored *Participation et Développement de l'Adolescence* program (in 56 schools and 22 other locations in six provinces), the *No SIDA* travelling caravan, various *Campus sans SIDA* campaigns, as well as school-based health clubs. All of these actions were conducted in a sporadic and haphazard manner, and lasted at most a few days. Individual initiatives on the part of teachers were the norm, and UNESCO reported in 2007 that there were isolated instances of secondary school teachers devoting one hour per month to HIV/AIDS related topics.

3.3 Program description

3.3.1 Teacher training on HIV/AIDS

The Cameroonian government, with the support of UNESCO, is currently in the process of introducing a *teaching module* devoted to HIV/AIDS into its school curriculum. The goal is to effectively educate schoolchildren concerning HIV/AIDS, by including information in the curriculum concerning prevention, transmission mechanisms, as well as other issues linked to HIV/AIDS. The official name of this new component of the Cameroonian school curriculum is "Education à la vie familiale, en matière de population et au VIH/SIDA", with "EVF /EMP/VIH et SIDA" being the most commonly-used acronym.

During the 2006-2007 academic year, roughly 2,000 teachers of French, English (the two official languages in the Cameroon), civic education, history and geography, life sciences (biology) and physical education participated in teacher training sessions on the HIV/AIDS teaching module in the five Cameroonian provinces of Adamaoua, Nord, Extrème Nord, Sud, and Centre.

Three options were envisaged by authorities in terms of HIV/AIDS education in the schools, and this choice ultimately determined, at least in part, the identity of the teachers who ended up participating. The first option was to introduce the "EVF/EMP/VIH et

SIDA" module as a subject in its own right, with the same status as English, French or mathematics. The second option was to introduce the module within a given discipline ("discipline d'accueil"), with the third option being to do so within several different disciplines. Ultimately, the third option was chosen, with the disciplines within which the module would be tought being (i) languages, (ii) civic education, (iii) history and geography, (iv) physical education and (v) life sciences.

In each participating school, nine teachers drawn from the five eligible disciplines were selected to participate in the training program. This choice may have been a function of the discretionary power of the school's director (and this is highly likely in the Cameroonian context), and may also have been driven by economic motives: the daily *per diem* granted to participants amounted to $100. For teachers who participated in the full three day program, participation was roughly equivalent to their monthly wage. The cost of the intervention was thus roughly $2,700 per treated school.

Within each province, training took place at 3 to 5 HIV/AIDS "focal points" of the Ministry of Secondary Education. More often than not, these corresponded to the equivalent of teacher training colleges ("Ecole Nationale des Instituteurs de l'Enseignement - ENIEG"), which are well attuned to this type of activity in that they have traditionally been associated with HIV/AIDS information campaigns in Cameroonian schools. The trainers, often, were inspectors ("inspecteurs pédagogiques nationaux") of the teacher's discipline, who had themselves been trained by UNESCO-provided external experts. Together, these external experts and the trainers formulated a pedagogical guide that served as a reference manual for the trainees. By and large, the content of teacher training was based on the logic model, in which trained teacher are tought (in theory) how to modify intermediate indicators so as to ultimately change student behavior. The intermediate student mediating factors focused on by the teacher training included knowledge, perceived risk, perceived severity, personal values and attitudes, perceived peer behavior and norms, self-efficacy and skills, motivation and intentions. Though it is difficult to pinpoint a particular psychological theory underlying the training, it is safe to say that it was based, at least in part, on the protective motivation, behavior change for intervention, theory of reasoned action, social learning and psychological determinants of behavior approaches.

Trained teachers are tought to define STDs and give examples of the most common ones, to speak of prevention methods, to describe the manners in which STDs manifest themselves, to describe their transmission mechanisms and the best prevention methods, to explain the socioeconomic consequences of STDs and, perhaps most importantly in light of the empirical results we present below, to teach students how to manipulate condoms. Trained teachers are also tought to define HIV/AIDS, to define a person's serological status, to promote voluntary testing, to explain that AIDS cannot be cured, to explain opportunistic infections in HIV-positive individuals, and to debunk various folk myths concerning the transmission and curing of HIV/AIDS. They are also incited to explain the propagation of the pandemic in the Cameroon, in Africa, and in the world, and to promote protective behavior.[4]

3.3.2 Selection of schools and program design

The choice of schools from which the potentially participating teachers were drawn was not random. Given that the program was funded by the African Development Bank (AfDB), participating schools were drawn from towns in which the AfDB had contributed to the construction or renovation, within the town's health center, of a unit explicitly devoted to HIV/AIDS prevention, information and testing.

The key criterion determining program deployment was the number of secondary schools in the town. For budgetary reasons, it was not possible to treat all schools in the country in which the AfDB had intervened at the health center level. On the other hand, the project management team wished to treat a relatively high proportion of schools in a given town, so as to avoid dispersing teacher training too thinly. The administrative rule that

[4] For details, see MINESEC (2006).

was eventually chosen was to treat towns in which between 1 and 4 secondary schools were present, and not treat the rest. Our sample is thus constituted by 108 schools in towns (i) where the AfDB had funded the construction or renovation of a town health center unit devoted to HIV/AIDS prevention, information and testing and (ii) in which there were between 1 and 8 schools. Roughly half the sample is constituted by schools in towns which fall below the 4 secondary school threshold, with the other half being in towns whose number of secondary schools is above the threshold. As we shall see below, this simple administrative rule that determined program deployment is the key to our identification strategy, and allows us to evaluate the impact of the intervention using a sharp regression discontinuity design.

3.4 Data and estimation strategy

3.4.1 The data

The sample used in this evaluation consists of 2,279 15-17 year olds (attending 2^{nde} and $1^{ère}$ classes) and 2,267 12-13 year olds (attending $6^{ème}$ and $5^{ème}$ classes) in Cameroonian towns in which there were between 1 and 8 secondary schools, and who were attending class between December 2007 and January 2008; 108 schools were involved; 56 schools were teacher training schools, whereas the remaining 52 schools constitute our control group. In each school, we surveyed four classes: one $6^{ème}$, one $5^{ème}$, one 2^{nde} and one $1^{ère}$ class.[5]

Power calculations for the 15-17 age category indicated that a sample of roughly this size, stratified over this number of schools, would be appropriate.[6] Such a sample should be of sufficient size, for example, to detect a 5% change in self-reported condom use, at the 95% level of confidence. In each school, between 10 and 12 pupils per class (divided equally between boys and girls) were randomly selected, as were the specific classes to be

[5] In the Cameroon, as in the old French school system, the numbering of classes works *backwards*.

[6] The power calculations were based on the 2004 Cameroon DHS, the 2001 Cameroon MICS and the 2000 ECAM II (Enquête Camerounaise Auprès des Ménages —basically a Cameroonian LSMS).

surveyed.

Our survey design for the lower age category could not be based on explicit power calculations, given the lack of information concerning HIV/AIDS-related issues for 12-13 year olds in existing statistics. As such, we applied the same sampling scheme to the lower age category as to the older students.

3.4.2 Teaching about HIV/AIDS

To determine how the teacher training intervention changed teaching about HIV/AIDS, we asked students whether HIV/AIDS was mentioned in class, how many teachers mentioned HIV/AIDS, and how many minutes per week were typically devoted by teachers mentioning HIV/AIDS. We also enquired as to the specific HIV/AIDS-related theme mentioned by these teachers, including: (i) reproductive health and STDs, (ii) identifying risky situations and vulnerability factors, (iii) developing everyday competencies to avoid risky situations, (iv) developing attitudes and competencies associated with safer sexual behavior, and (v) participating in the fight against HIV/AIDS.

3.4.3 Knowledge, attitudes, and self-reported behavior

A Knowledge, Attitude and Behavior (KAB) survey was conducted amongst the sampled students. The survey was constituted by a self-administered questionnaire in French (the language of instruction in all of the schools in our sample). Given the timing of teacher training during the 2006-2007 academic year, the survey was conducted approximately 10 months after teacher training occurred.

Specific questionnaires were prepared for the two age brackets. For the younger students, the questionnaire focused on the KAB themes, with a particular emphasis on the first two themes. Specific questions involve knowledge concerning transmission mechanisms and protection, attitudes towards infected persons and stigmatisation. For the 15-17 year age bracket, a slightly different questionnaire was administered, given that more emphasis is placed on sexual behavior. While collecting information on the KAB themes, we also collected information on each student's family background. We also administered a separate questionnaires to all teachers, whether they followed the teacher training program or not.

For the 12-13 year olds, the key response variables, divided by KAB theme, include:

- *knowledge*: student knows that HIV causes AIDS, knows that mosquito bites do not transmit HIV, knows that having a STD increases the likelihood of infection;
- *attitudes*: student would care for a relative who has HIV/AIDS, would buy food from a shopkeeper who has HIV/AIDS, and declares that she/he will remain sexually abstinent during adolescence;
- *behavior*: student declares that he/she has never had sex, that he/she has had more than one partner, and that he/she used a condom the last time he/she had sex.

For the 15-17 year olds, the response variables are similar; with minor variations; in particular, we consider whether the student has contracted a sexually-transmitted disease during the last 12 months; and whether she/he has subjected him/herself to an HIV/AIDS test during the last 12 months

.

3.4.4 Childbearing

In administering the KAB questionnaire to 15-17 year old girls, we also collected information on childbearing, pregnancy and marital status. As noted by Duflo, Dupas, Kremer, and Sinei (2006) and Dupas (2009), self-reported behavior may be contaminated by reporting bias, and child bearing is an objective proxy for unprotected sex, though an imperfect one. As such, and as in the studies just mentioned, our main measure of whether attending a teacher training school changed the probability that students engaged in risky sexual behavior is the incidence of childbearing among girls

.

3.4.5 Identification strategy

Our identification strategy exploits the natural experiment generated by being either just above or just below the 4 secondary schools per town threshold adopted as the key determinant of program deployment. This sharp regression discontinuity design (RDD) approach is akin to a randomization which is valid around the threshold: in the words of Lee and Lemieux (2010), it is a "local randomized experiment". An RDD identification strategy is valid if individuals are unable to precisely control the forcing variable (i.e. the variable whose threshold value determines treatment status). Since the number of secondary schools in a Cameroonian town is determined by the town's population and administrative decisions taken several years earlier at the Ministry of Secondary Education, it is difficult to see how it could have been manipulated by interested parties so as to induce inclusion of schools in the teacher training program.

RDD was introduced into the economics literature in a widely-cited paper on the relationship between class size and academic performance by Angrist and Lavy (1999). A formal econometric treatment is given in Hahn, Todd, and Van Der Klaauw (2001) and, for the case of a discrete forcing variable, in Lee and Card (2008). Excellent surveys have recently been provided by Imbens and Lemieux (2008) and Lee and Lemieux (2010).

Suppose that we have a forcing variable denoted by X_s, where $s = 1, ..., S$ indexes schools, define:

$$D_s = \mathbf{1}[X_s > c],$$

where $\mathbf{1}[.]$ is the indicator function and c will denote the threshold of 4 secondary schools in the town. Letting Y_{is} denote the response variable of interest for student $i = 1, ...N$ in school s, with $Y_{is}(1)$ denoting treated schools and $Y_{is}(0)$ denoting untreated schools, average potential outcomes, evaluated at the threshold c, can be written as:

$$E[Y_{is}(1)\,|X_s = c] \text{ and } E[Y_{is}(0)\,|X_s = c].$$

Intuitively, what we would like to estimate is:

$$\lim_{X \downarrow c} E[Y_{is}\,|X_s = c] - \lim_{X \uparrow c} E[Y_{is}\,|X_s = c]$$

which corresponds to:

$$E[Y_{is}(1) - Y_{is}(0)\,|X_s = c].$$

In practice, this is implemented in a simple regression framework. In terms of potential outcomes, write (Lee and Lemieux (2010), p. 37):

$$Y_{is}^L = \alpha_L + (X_s - c)\beta_L + \varepsilon_{is}$$

for observations to the left of the threshold, and

$$Y_{is}^R = \alpha_R + (X_s - c)\beta_R + \varepsilon_{is},$$

for observations on its right, where β_L and β_R are the regression coefficients associated with the forcing variable (to the left and to the right of the threshold, respectively). One can combine the two potential outcomes by posing:

$$Y_{is} = Y_{is}^L(1 - D_s) + Y_{is}^R D_s,$$

which implies that the simple linear regression being estimated is given by:

$$Y_{is} = \alpha_L + D_s\,(\alpha_R - \alpha_L) + D_s(X_s - c)\beta_R + (1 - D_s)(X_s - c)\beta_L + \varepsilon_{is}. \tag{3.1}$$

The RDD estimate of the program's impact is then given by the coefficient $(\alpha_R - \alpha_L)$ associated with the treatment dummy in (3.1), and allows for different slopes in terms of the relationship between the forcing variable and the outcome to the left and to the right of the threshold.

In terms of inference, we assume that disturbance terms are independent across schools, but are allowed to be correlated across observations within the same school. We also performed inference while clustering standard errors on the basis of the discrete values taken on by the forcing variable X_s (the number of secondary schools in the town), as suggested in a recent paper by Lee and Card (2008). In this case, standard errors were significantly lower than those reported here. We also combined both approaches to inference by implementing two-way clustering Cameron, Jonah B., and Miller (2006) in which the disturbance terms were allowed to be correlated both across observations within the same school, and across observations corresponding to a given value of the forcing variable. Again, this resulted in slightly more precise estimates. In order to err on the side of caution, we report standard errors clustered solely at the school level in all tables.

We also ran regressions as in (3.1) with various individual (student, parent) and school covariates included. Given the quasi-experimental nature of our RDD-based identification strategy, the results are almost identical, so we restrict our reporting of results to those based on the specification given in (3.1).[7]

In the results reported below, we consider two samples of schools. Our full sample is constituted by students in all schools located in towns with between 1 and 8 secondary schools. Our discontinuity sample restricts its attention to schools located in towns with between 3 and 6 secondary schools. To the left of the threshold, there are therefore students in schools from towns with either 3 or 4 secondary schools. To the right of the threshold,

[7]One can also estimate (3.1) while including a polynomial in X_s. Given the relatively small number of values that can be taken by the number of schools in the town (from 1 to 4 on the LHS of threshold, and from 5 to 8 on the RHS), there is little to be gained from doing so. The same argument applies to estimating (3.1) semiparametrically. Any reasonable bandwidth choice results in the linear specification given by (3.1).

there are student in schools from towns with either 5 or (sometimes) 6 secondary schools. The results based on the discontinuity sample constitute our preferred specification, in that they provide a better approximation to a local randomized experiment. The downside is that some statistical power is lost in that the sample is significantly smaller. Figures 3.1 and 3.2 provide histograms of the distribution of students in our sample (for each age bracket), according to the number of secondary schools present in the town.

Table 3.1 presents the observable characteristics of the schools and teachers included in our samples[8]. These school characteristics should not be affected by the teacher training status of the school and should be statistically indistinguishable between treatment and control schools if our RDD identification strategy is indeed akin to a local randomization.

As should be obvious from the results, this is the case for the full sample, with the exception of the number of classes tought by each teacher and the number of students tought, with teachers in control group schools teaching less classes and less students. If anything, this should bias our results against finding any statistically significant effect of school teacher training status, since the quality of education may be slightly better in control schools. For the discontinuity sample, there is no statistically significant difference between teacher training and control schools at the 5% level of confidence, and no $p-$values are lower than 0.072.

Table 3.2 presents similar descriptive statistics for students (and the households that they belong to), for both age categories and for both the full and discontinuity samples. With the exception of whether the household head works (for the 15-17 year age category) none of the student and student household characteristics differ significantly between the treated and control samples.[9]

[8]Our data are confined to the post-treatment period. Given that the RDD identification strategy is akin to a local randomization, baseline (pre-intervention) data are required. We consider various 'placebo' RDD regressions below in order to assess the robustness of our identification strategy.

[9]It is also worth noting that the mean age of students in the 15-17 year age category is actually approximately 18: this is common in the Cameroonian school system where many students either suffer from spells out of school (for example, when the household head loses his or her job) or repeat grades.

3.5 Results

3.5.1 Delivery of HIV/AIDS education

Table 3.3 presents RDD results, based on the specification given in equation (3.1), of how teaching about HIV/AIDS was affected by teacher training. In contrast to the results reported by Duflo, Dupas, Kremer, and Sinei (2006) for the Kenyan case —in which all measures of teaching HIV/AIDS were higher in teacher training schools— the Cameroonian teacher training program resulted in a significant shift in focus of teaching HIV/AIDS related issues. First, less teacher time was devoted to the 12-13 year old category, with the time devoted to the 15-17 year olds remaining roughly constant. Second, the message being delivered to both age categories changed: away from general issues relating to reproductive health and the fight against HIV/AIDS, and towards practical knowledge in terms of avoiding risky behavior and adopting safe-sex techniques.

The first shift in focus is apparent in three measures of the importance accorded to HIV/AID by the teachers of 12-13 year olds. First, the likelihood that at least one teacher mentions HIV in class is significantly lower for this age category in teacher training schools than in control schools (at least for our preferred discontinuity sample results). Second, the number of teachers reported by students to be mentioning HIV/AIDS in class is significantly lower for 12-13 year olds in teacher training schools than in control schools. Third, for teachers who do mention HIV/AIDS in class, the time devoted to the topic is significantly lower for 12-13 year old students attending teacher training schools. This result is robust to estimating the equation in question by Poisson regression or zero-inflated Poisson (ZIP).

The second, and in all likelihood the most important, manner in which teacher training has affected classroom behavior involves the themes addressed by teachers when they mention HIV/AIDS. In teacher training schools, teachers (based on the discontinuity sample results) are significantly less likely to discuss reproductive health and STDs, or risky situations and vulnerability factors with 15-17 year olds, but much more likely to discuss attitudes and competencies for safer sexual behavior. With 12-13 year olds they are much more likely to discuss developing everyday competencies to avoid risky behavior, and much less likely to discuss participating in the fight against HIV. All of these results are qualitatively robust to rephrasing things in terms of a multinomial logit specification. Thus, the emphasis in teacher training schools has shifted to everyday practices that should lead to safer sex, at the expense of general knowledge concerning HIV/AIDS and STDs. For 15-17 year olds, the quantity of HIV/AIDS related information has arguably remained the same, but its practical quality has improved. For 12-13 year olds, the quantity has decreased, but the quality has arguably increased as for the 15-17 year olds. For the 15-17 year olds, the net result should be positive in terms of behavior. For the 12-13 year olds, the outcome in terms of behavior depends upon the quantity/quality tradeoff. Note also that if this shift in emphasis leads to changes in student KAB, one would expect to see some indication of this in safer sexual practices on the one hand, but worse knowledge, on the other.

The upshot is that HIV/AIDS teacher training in the Cameroon appears to have resulted in (i) a quantitative reduction of HIV/AIDS teaching resources devoted to the younger age category, and to have resulted (ii) in greater emphasis being placed, *for both age categories*, on prevention and safe sex practices and less emphasis on knowledge concerning HIV/AIDS and STDs in general. As we shall see below, this shift in emphasis towards safe-sex practices is indeed reflected in outcomes, both in terms of self-reported sexual knowledge and behavior and, more importantly, in terms of teenage childbearing.

3.5.2 Knowledge, attitudes, and self-reported behavior

Tables 3.4 and 3.5 presents RDD results, again based on the specification given in equation (3.1), of how students in teacher training schools differ from their brethren in control schools, with respect to self-reported knowledge, attitudes, and behavior. For 15-17 year olds, two results stand out.

First, older students are significantly less likely to know that a person in apparent good health can be HIV-positive. The result holds for boys and girls, in both the full and discontinuity samples. Boys are also significantly less likely to believe that abstinence is the main way of avoiding contamination (again, this is true in both samples).

Second, in what might be a direct consequence of the previous result, both boys and girls in teacher training schools are significantly more likely to have used a condom during their last sexual intercourse. This is exactly what one would expect, given the increased emphasis on safe sexual practices in terms of what is being tought in the classrooms of teacher training schools.

Note that these results stand in direct contrast to two results reported for Kenya by Duflo, Dupas, Kremer, and Sinei (2006): boys were more likely to mention abstinence a means of preventing HIV infection, and were not more likely to have reported using a condom during their last sexual encounter. These differences between the Cameroon and Kenya highlight that the specific content of HIV/AIDS instruction being dispensed in teacher training schools matters: in the Kenyan teacher training program, teachers were *not* trained to promote condoms. But in schools in which condom debates were organized (a separate intervention that is evaluated in Duflo, Dupas, Kremer, and Sinei (2006) alongside teacher training and subsidizing school uniforms), the likelihood of having used a condom during one's last sexual encounter was indeed higher. It would therefore appear, at least from the Cameroonian and Kenyan experiences, that self-reported condom use can indeed be significantly increased either by having trained teachers focus on safer sexual practices and condom manipulation in class (as in the Cameroon), or by organizing condom debates and essay competitions (as in Kenya). In the Cameroonian case, because of the shift in emphasis in class away from general knowledge concerning HIV/AIDS and STDs, this comes at the expense of worse self-reported knowledge outcomes.

For the 12-13 year olds, the results reported in Table 3.5 confirm the impact of the within-school reallocation of HIV/AIDS teaching time, as well as the choice of topics covered. First, although children in teacher training schools are in general more likely to know that HIV causes AIDS, they are significantly less likely to know that STDs increase risk (this result holds for girls in both samples, and for boys in the full sample). Girls, at least based on the discontinuity sample results, are also significantly less likely to be willing to buy food from a shopkeeper who is HIV-positive (in contrast, in Kenya girls became more tolerant in teacher training schools).

Second, as with the 15-17 year olds, and despite the reduction in teaching time devoted to HIV/AIDS topics (but in all likelihood because of the shift in emphasis by teachers towards developing everyday competencies to reduce risky behavior), 12-13 year olds are (based on the discontinuity sample results) significantly more likely in teacher training schools to have never had sexual intercourse and to have used a condom if they have had sexual intercourse, and are significantly less likely to have had sex with more than one partner[10].

3.5.3 Childbearing rates

Table 3.6 presents RDD estimates of the effect of teacher training on childbearing rates for 15-17 year old girls[11]. In contrast to the results reported in Duflo, Dupas, Kremer, and Sinei (2006) for Kenya, Cameroonian girls are significantly less likely to have begun childbearing if they attend a teacher training school. For the full sample, the reported coefficient indicates a 7 percentage point decrease in the incidence of childbearing for girls in teacher training schools with respect to girls in control schools. For the discontinuity sample, the corresponding number is 10 percentage points.

Combined with the earlier findings that 15-17 year old girls are significantly more likely

[10] Note that there may also be spillover effects from the older to the younger students. On the basis of our fieldwork, our conjecture would be that this is not the case in the Cameroonian context, though there is no manner of being certain.

[11] Note that, since we do not have baseline data, we cannot be entirely sure that teenage pregnancy rates were not higher in control schools ex ante, though there is no a priori to think that this might have been the case. On the other hand, as reported below, placebo regressions in which we replace the true hreshold value of the forcing variable by other values fail to turn up any statistically significant difference in teenage childbearing rates. Our result on the impact of teacher training on teenage childbearing therefore only obtains for the true threshold value of the forcing variable, thereby lending it additional credence.

to have used a condom during their last sexual intercourse and that teachers in teacher training schools are significantly more likely to place more emphasis on developing attitudes and competencies for safer sexual behavior (as well as devoting significantly more class time to HIV/AIDS topics) suggests that teacher training may indeed have an impact on HIV transmission rates for teenage girls.

Note also, in contrast to Duflo, Dupas, Kremer, and Sinei (2006), that the marriage rate has not increased, and that there is no evidence that girls who have started childbearing are more likely to be married to the father of their child in teacher training schools. In the Cameroonian context, this is not surprizing: very few of the 15-17 year old girls in our sample are married. Indeed, the results concerning childbearing are virtually identical if one restricts one's attention to unmarried students. Though we do not have information, such as that gathered in Dupas (2009), concerning the age of the partners of girls who have started childbearing, we have found no significant change, for girls in teacher training schools, in terms of the self-reported likelihood of their having had sex, within the past 12 months, with a partner to whom they are not married and with whom they are not living.[12] This suggests that girls in teacher training schools have indeed significantly increased condom use, and that this has resulted in a significant reduction in the incidence of childbearing. To the extent that childbearing is a valid proxy for risky sexual behavior, this suggests that teacher training may have significantly reduced HIV transmission rates, though confirmation of this result must await data which encompass explicit biomarkers.

3.5.4 Robustness

In order to check whether our identification strategy was robust, we implemented a plethora of "placebo" RDD regressions, in which the threshold number of schools in the town was set at anothervalue, different from the one actually associated with the determination of school treatment status. In Table 3.7, we report a selection of these, corresponding to our results

[12] In the interests of brevity we do not report this result explicitly. It is, of course, available upon request.

for self-reported behavior and pregnancy for the older age category. If our identification strategy based on the discontinuity induced by the four schools per town threshold were invalid, one would expect to find unwarranted significant effects at other "threshold" values that do not actually correspond to the true distinction between treated and control schools.

As shown by the results reported in Table 3.7, in which we replace the true treatment dummy of Equation 3.1 with a placebo treatment dummy which is equal to one when the town has five or less schools and is equal to zero otherwise, this is not the case[13]. For teenage childbearing, be it in the full or in the discontinuity sample, the 'impact' of the placebo treatment dummy is statistically indistinguishable from zero, whereas it was highly significant when using the "true" treatment dummy, as reported in Table 3.6. Similarly, the placebo treatment effect estimated for condom use is statistically insignificant for all of the results reported in Table 3.7 ,whereas the corresponding coefficients were highly significant in three out of four cases in the results reported earlier in Table 3.6 . The same is true for the results concerning whether the student carried out an HIV test.

Though these placebo regression results cannot prove that our identification strategy is valid, they do strongly suggest that the statistically significant impact of teacher training on teenage childbearing and on self-reported behavior that we uncovered above is causal, and is not driven by a spurious correlation.

3.6 Conclusion

Should governments be investing in HIV/AIDS teacher training programs in Sub-Saharan Africa, such as the one considered in this chapter, and as they are currently doing? Though it is well known that targeted interventions amongst high-risk populations (such as sex workers) are effective in reducing the spread of HIV/AIDS, interventions in large, relatively low risk populations are often seen as being impractical and prohibitively costly. On the

[13] Results are qualitatively similar at other placebo threshold values of the forcing variable, with no statistically significant impact of teacher training on teenage childbearing.

other hand, it would appear to be especially important to consider interventions in such populations, in that significant increases in HIV infection will necessarily come, in the future, from the lower risk general population.

Given the near-universality of schooling in relatively rich African countries such as the Cameroon, our results on teenage childbearing and self-reported condom use for 15-17 year old girls suggest that such programs may be an effective manner of reaching an important segment of the general population (this may not be the case in much poorer African countries, where enrollment rates are much lower). Moreover, for the Cameroonian case, in which prevalence rates are significantly higher than in most of francophone West Africa, it would appear to be extremely important to implement any intervention that can contribute to maintaining prevalence below the threshold of 5%.

Though we do not have quantitative evidence on the effectiveness of alternative interventions in the Cameroon, such as blood screening, the use of the mass media, the social marketing of condoms, the treatment of STDs, the peer education of sex workers, voluntary counselling and testing, prevention activities among intravenous drug users or the prevention of mother-to-child transmission, our empirical results suggest that teacher training can have a significant impact on the behavior of 15-17 year olds. Whether this is cost-effective, however, is another matter. Given the relatively high cost of training teachers ($US 2,700 per school) and on the basis of the school uniform and condom debate results reported for Kenya by Duflo, Dupas, Kremer, and Sinei (2006), this might not be the case.

Our results also highlight an interesting within-school quantity-quality tradeoff in terms of treatment. For the 15-17 year olds, teacher training involved both an increase in quantity (class time) and a change in quality (a shift towards more emphasis on safe sex practices) of teaching regarding HIV/AIDS. The result, in contrast to the Kenyan case, is that girls in teacher training schools in the Cameroon are significantly less likely to have begun childbearing, probably because of a significant increase in condom use. For the 12-13 year olds, on the other hand, the negative effect of the decrease in quantity (class time) appears to have been trumped by the change in quality (more emphasis on developing everyday competencies to avoid risky behavior), with the net impact of these two contrasting effects being a significant improvement in self-reported sexual behavior.

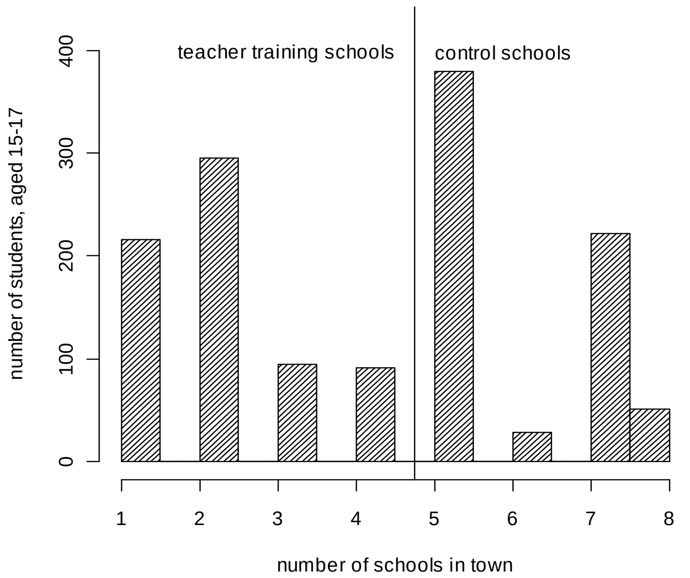


Figure 3.1: The distribution of students, aged 15-17. Teacher training schools are those in towns with between 1 and 4 secondary schools. Control schools are those in towns with between 5 and 8 secondary schools.

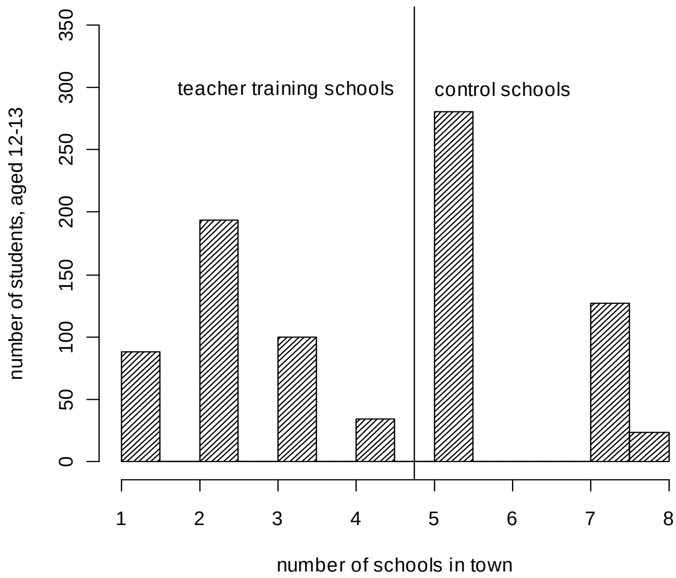


Figure 3.2: The distribution of students, aged 12-13. Teacher training schools are those in towns with between 1 and 4 secondary schools. Control schools are those in towns with between 5 and 8 secondary schools.

	Teacher training schools: $D = 1$	Control schools: $D = 0$	Treatment − control	p−value
	Full sample 1 to 8 secondary schools in town			
Teacher characteristics				
age (in years)	35.438	35.257	0.180	0.836
sex (male = 1)	0.790	0.779	0.010	0.816
experience (in years)	8.631	8.395	0.236	0.771
number of classes tought	6.044	4.719	1.324	0.000
number of students tought	296.953	247.212	49.741	0.015
School characteristics				
has an infirmary	0.328	0.437	−0.108	0.098
has electricity	0.749	0.856	−0.106	0.063
has toilets	0.875	0.898	−0.022	0.575
has water	0.616	0.696	−0.079	0.192
students per class	62.677	58.242	−4.435	0.174
Number of schools	56	52		
	Discontinuity sample 3 to 6 secondary schools in town			
Teacher characteristics				
age (in years)	34.729	36.823	−2.094	0.072
sex (male = 1)	0.723	0.730	−0.007	0.929
experience (in years)	8.374	9.355	−0.981	0.396
number of classes tought	6.340	6.041	0.299	0.754
number of students tought	268.495	220.602	47.893	0.152
School characteristics				
has an infirmary	0.532	0.401	0.130	0.199
has electricity	0.833	0.837	−0.004	0.962
has toilets	0.866	0.893	−0.040	0.740
has water	0.687	0.614	0.072	0.437
students per class	56.391	55.075	1.315	0.806
Number of schools	18	32		

Table 3.1: Observable characteristics of teachers and schools in the full and discontinuity samples.

	Teacher training schools: $D=1$	Control schools: $D=0$	Treatment − control	p−value
	Full sample 1 to 8 secondary schools in town			
Student characteristics: 15-17 year olds				
age (in years)	18.273	18.146	0.127	0.483
sex (male = 1)	0.533	0.525	0.008	0.656
household size	7.675	7.763	−0.087	0.835
household head works	0.763	0.773	−0.009	0.743
at least one parent can read	0.941	0.948	−0.006	0.696
Number of students	1,273	1,248		
Student characteristics: 12-13 year olds				
age (in years)	13.575	13.480	0.094	0.603
sex (male = 1)	0.496	0.515	−0.018	0.203
household size	8.203	8.290	−0.087	0.774
household head works	0.863	0.835	0.028	0.148
at least one parent can read	0.939	0.951	−0.012	0.489
Number of students	1,337	1,304		
	Discontinuity sample 3 to 6 secondary schools in town			
Student characteristics: 15-17 year olds				
age (in years)	18.498	18.405	0.093	0.723
sex (male = 1)	0.526	0.514	0.011	0.692
household size	7.545	6.889	0.656	0.134
household head works	0.781	0.708	0.073	0.070
at least one parent can read	0.942	0.974	−0.031	0.143
Number of students	403	733		
Student characteristics: 12-13 year olds				
age (in years)	13.731	13.520	0.211	0.468
sex (male = 1)	0.506	0.515	−0.008	0.761
household size	7.997	7.725	0.272	0.437
household head works	0.822	0.823	−0.001	0.969
at least one parent can read	0.968	0.969	−0.001	0.938
Number of students	444	798		

Table 3.2: Observable characteristics of students in the full and discontinuity samples.

	RDD estimates of the impact of teacher training			
	Full sample		Discontinuity sample	
Teacher mentions HIV in class				
15-17 year olds	−0.006 (0.095)	$N = 2,279$ *mean*=0.845	−0.094 (0.195)	$N = 1,120$ *mean*=0.866
12-13 year olds	−0.100 (0.106)	$N = 2,667$ *mean*=0.333	−0.594 (0.128)	$N = 1,249$ *mean*=0.345
Number of teachers teaching HIV				
15-17 year olds	−0.290 (0.247)	$N = 2,577$ *mean*=1.046	−0.022 (0.906)	$N = 1,162$ *mean*=0.957
12-13 year olds	−0.108 (0.080)	$N = 2,667$ *mean*=0.209	−0.430 (0.098)	$N = 1,249$ *mean*=0.231
Weekly class time (minutes) devoted to HIV				
15-17 year olds	2.611 (5.117)	$N = 2,577$ *mean*=20.33	8.124 (17.191)	$N = 1,162$ *mean*=17.44
12-13 year olds	−3.452 (4.875)	$N = 2,667$ *mean*=12.90	−25.388 (5.386)	$N = 1,249$ *mean*=13.16
For teachers who mention HIV/AIDS: specific theme mentioned				
developing everyday competencies to avoid risky behavior				
15-17 year olds	0.134 (0.058)	$N = 1,380$ *mean*=0.039	0.171 (0.085)	$N = 596$ *mean*=0.068
12-13 year olds	0.133 (0.101)	$N = 846$ *mean*=0.042	0.753 (0.365)	$N = 415$ *mean*=0.072
attitudes and competencies for safer sexual behavior				
15-17 year olds	0.173 (0.078)	$N = 1,380$ *mean*=0.059	0.293 (0.114)	$N = 596$ *mean*=0.072
12-13 year olds	0.005 (0.051)	$N = 846$ *mean*=0.056	0.007 (0.057)	$N = 415$ *mean*=0.024
participating in the fight against HIV				
15-17 year olds	−0.238 (0.127)	$N = 1,380$ *mean*=0.302	0.146 (0.137)	$N = 596$ *mean*=0.344
12-13 year olds	−0.158 (0.155)	$N = 846$ *mean*=0.260	−0.668 (0.148)	$N = 415$ *mean*=0.304
reproductive health and STDs				
15-17 year olds	−0.057 (0.148)	$N = 1,380$ *mean*=0.399	−0.387 (0.191)	$N = 596$ *mean*=0.347
12-13 year olds	−0.012 (0.164)	$N = 846$ *mean*=0.401	0.340 (0.296)	$N = 415$ *mean*=0.340

Table 3.3: How different are things in the classroom in teacher training schools? Full sample includes all schools in towns with between 1 and 8 schools. Discontinuity sample is restricted to schools in towns with between 3 and 6 schools. Standard errors clustered at the school level in parentheses.

	RDD estimate of the impact of teacher training			
	Full sample		Discontinuity sample	
	Boys	Girls	Boys	Girls
15-17 year olds: Response variables				
Knowledge:				
Person apparently in	-0.059 (0.029)	-0.117 (0.041)	-0.223 (0.111)	-0.367 (0.214)
good health can be HIV+	$N = 1,351$ *mean*=0.927	$N = 1,190$ *mean*=0.932	$N = 600$ *mean*=0.930	$N = 550$ *mean*=0.930
Abstinence main way	-0.194 (0.081)	-0.002 (0.087)	-0.236 (0.128)	-0.106 (0.131)
to avoid contamination	$N = 1,329$ *mean*=0.522	$N = 1,173$ *mean*=0.528	$N = 588$ *mean*=0.542	$N = 535$ *mean*=0.527
Contraception is	0.018 (0.067)	0.045 (0.064)	0.126 (0.106)	-0.025 (0.089)
different from prevention	$N = 1,357$ *mean*=0.6205	$N = 1,203$ *mean*=0.670	$N = 595$ *mean*=0.626	$N = 553$ *mean*=0.676
Attitudes:				
Ready to meet and	0.052 (0.041)	0.028 (0.035)	-0.020 (0.118)	-0.178 (0.168)
assist HIV+ individuals	$N = 1,363$ *mean*=0.893	$N = 1,205$ *mean*=0.897	$N = 601$ *mean*=0.888	$N = 555$ *mean*=0.893
Discusses HIV/AIDS	-0.083 (0.068)	0.000 (0.051)	-0.023 (0.143)	-0.147 (0.076)
with family members	$N = 1,367$ *mean*=0.709	$N = 1,204$ *mean*=0.776	$N = 605$ *mean*=0.732	$N = 555$ *mean*=0.791
Worries that blade is	0.146 (0.071)	0.188 (0.081)	0.134 (0.086)	-0.062 (0.162)
single-use at barbershop	$N = 1,336$ *mean*=0.631	$N = 1,014$ *mean*=0.646	$N = 594$ *mean*=0.557	$N = 452$ *mean*=0.601
Behavior:				
Used a condom during	0.128 (0.081)	0.247 (0.092)	0.553 (0.110)	0.526 (0.221)
last sexual intercourse	$N = 982$ *mean*=0.590	$N = 692$ *mean*=0.604	$N = 469$ *mean*=0.607	$N = 355$ *mean*=0.566
Contracted a STD	-0.036 (0.055)	-0.053 (0.045)	0.033 (0.080)	0.009 (0.129)
during the past 12 months	$N = 870$ *mean*=0.109	$N = 634$ *mean*=0.083	$N = 424$ *mean*=0.115	$N = 329$ *mean*=0.066
Carried out	0.164 (0.066)	0.198 (0.076)	0.049 (0.129)	0.104 (0.151)
an HIV test	$N = 1,367$ *mean*=0.296	$N = 1,205$ *mean*=0.351	$N = 604$ *mean*=0.284	$N = 555$ *mean*=0.382

Table 3.4: RDD estimates of the impact of teacher training on KAB of 15-17 year olds. Full sample includes all schools in towns with between 1 and 8 schools. Discontinuity sample is restricted to schools in towns with between 3 and 6 schools. Standard errors clustered at the school level in parentheses.

	RDD estimate of the impact of teacher training			
	Full sample		Discontinuity sample	
	Boys	Girls	Boys	Girls
12-13 year olds: Response variables				
Knowledge:				
HIV	0.158	0.236	0.342	0.226
	(0.071)	(0.084)	(0.123)	(0.162)
causes AIDS	$N = 1,330$	$N = 1,297$	$N = 627$	$N = 604$
	$mean$=0.409	$mean$=0.400	$mean$=0.390	$mean$=0.346
Mosquito does	0.037	0.083	−0.202	−0.145
	(0.080)	(0.084)	(0.138)	(0.123)
not transmit AIDS	$N = 1,349$	$N = 1,313$	$N = 638$	$N = 608$
	$mean$=0.415	$mean$=0.358	$mean$=0.366	$mean$=0.315
STDs increase	0.127	−0.126	0.083	−0.436
	(0.072)	(0.068)	(0.135)	(0.116)
risk	$N = 1,325$	$N = 1,284$	$N = 673$	$N = 595$
	$mean$=0.518	$mean$=0.541	$mean$=0.508	$mean$=0.529
Attitudes:				
Willing to take care	0.022	−0.0187	−0.106	−0.104
	(0.042)	(0.038)	(0.140)	(0.097)
of HIV+ family member	$N = 1,347$	$N = 1,312$	$N = 639$	$N = 610$
	$mean$=0.842	$mean$=0.866	$mean$=0.865	$mean$=0.882
Willing to buy from	−0.042	0.014	−0.397	0.058
	(0.078)	(0.070)	(0.104)	(0.123)
shopkeeper who is HIV+	$N = 1,325$	$N = 1,304$	$N = 627$	$N = 609$
	$mean$=0.514	$mean$=0.424	$mean$=0.491	$mean$=0.403
Will remain abstinent	0.160	0.0724	0.121	0.047
	(0.090)	(0.103)	(0.217)	(0.155)
during adolescence	$N = 788$	$N = 939$	$N = 362$	$N = 457$
	$mean$=0.647	$mean$=0.697	$mean$=0.610	$mean$=0.669
Behavior:				
Has never had	0.135	0.008	0.255	0.266
	(0.098)	(0.089)	(0.159)	(0.103)
sexual intercourse	$N = 1,335$	$N = 1,305$	$N = 633$	$N = 606$
	$mean$=0.591	$mean$=0.724	$mean$=0.571	$mean$=0.755
Has sex with more	−0.142*	−0.044	−0.205	−0.126
	(0.084)	(0.040)	(0.144)	(0.052)
than one partner	$N = 1,335$	$N = 1,305$	$N = 633$	$N = 604$
	$mean$=0.215	$mean$=0.0812	$mean$=−0.265	$mean$=0.094
Used a condom if	0.020	0.058	0.236	0.216
	(0.121)	(0.140)	(0.204)	(0.087)
had sexual intercourse	$N = 426$	$N = 198$	$N = 242$	$N = 106$
	$mean$=0.570	$mean$=0.712	$mean$=0.562	$mean$=0.790

Table 3.5: RDD estimates of the impact of teacher training on KAB of 12-13 year olds. Full sample includes all schools in towns with between 1 and 8 schools. Discontinuity sample is restricted to schools in towns with between 3 and 6 schools. Standard errors clustered at the school level in parentheses.

	RDD estimate of the impact of teacher training	
	Full sample Girls	Discontinuity sample Girls
Has started childbearing:	-0.077 (0.042)	-0.105 (0.045)
	$N = 1,144$ *mean*=0.118	$N = 702$ *mean*=0.102
Is married	-0.023 (0.025)	-0.019 (0.026)
	$N = 1,206$ *mean*=0.032	$N = 735$ *mean*=0.021
If has started childbearing:	-0.007 (0.009)	-0.0006 (0.0007)
Is married	$N = 1,144$ *mean*=0.012	$N = 702$ *mean*=0.004

Table 3.6: RDD estimates of the impact of teacher training on childbearing and marriage for 15-17 year old girls. Full sample includes all schools in towns with between 1 and 8 schools. Discontinuity sample is restricted to schools in towns with between 3 and 6 schools. Standard errors clustered at the school level in parentheses.

	Placebo RDD estimates of the impact of teacher training			
	Full sample		Discontinuity sample	
	Boys	Girls	Boys	Girls
Has started childbearing		−0.068 (0.048) $N = 1,144$ *mean*=0.118		−0.104 (0.083) $N = 702$ *mean*=0.102
Self-reported behavior:				
Used a condom during last sexual intercourse	0.069 (0.081) $N = 982$ *mean*=0.590	0.018 (0.136) $N = 692$ *mean*=0.604	0.036 (0.090) $N = 469$ *mean*=0.607	0.017 (0.143) $N = 355$ *mean*=0.566
Carried out an HIV test	0.000 (0.004) $N = 1,367$ *mean*=0.296	0.004 (0.079) $N = 1,205$ *mean*=0.351	−0.042 (0.073) $N = 604$ *mean*=0.284	−0.124 (0.089) $N = 555$ *mean*=0.382

Table 3.7: Placebo RDD regressions of the impact of teacher training on child-bearing and self-reported behavior of 15-17 year olds. "True" treatment cutoff of forcing variable: towns with 4 or less schools. Placebo cutoff value of forcing variable: towns with 5 or less schools. Full sample includes all schools in towns with between 1 and 8 schools. Discontinuity sample is restricted to schools in towns with between 3 and 6 schools. Standard errors clustered at the school level in parentheses.

General Conclusion

Civil war and HIV/AIDS in sub-Saharan Africa decimate the human capital necessary for growth and economic development. While civil war carries short and long term effects including loss of life, destruction of infrastructure of all kinds, destruction of social capital, disruption of institutional functioning, and the rolling back of progress made by countries in their development process, HIV/AIDS affects the productive capacity of young people and their households. These drivers of morbidity compromise the work force and affect resource allocation in households. When a HIV/AIDS positive household member becomes ill, the household resources are used to buy drugs, therefore reducing resources devoted to increased production and child education, augmenting the possibility for intergenerational transfer of poverty. Additionally, HIV orphans can be victims of malnutrition and lack of schooling, presenting a direct effect on the human capital of present and future generations. At the government level, HIV/AIDS also entails a change in the allocation of resources that could otherwise be used in the development process.

Despite several studies on the determinants of civil war and HIV/AIDS in Sub-Saharan Africa, very few studies are able to show a causal relationship at both macroeconomic and microeconomic level between civil war, HIV/AIDS and causal determinants. Failing to act directly on the determinants of civil war and HIV/AIDS it could be useful to act downstream. Specifically, regarding civil war it is important to understand how civil war affects households, who are the most affected group, which is the magnitude of negative impact of civil war. The answers to these questions are essential to formulate appropriate strategies to mitigate the effects of civil war on the human capital. Regarding HIV/AIDS it is important to find the most effective way to reach and eradicate the contamination among population group who are most affected, in sub-Sahara Africa namely the young population.

In this thesis we asked three main questions – namely, what are the short-term and long-term impacts of civil war in Angola on human capital? How effective are social programs during conflicts? And how is the effectiveness of social programs affected by conflict intensity? Does teacher training on HIV/AIDS have an impact on the level of knowledge, attitudes, and behavior related to HIV/AIDS in Cameroon?

In the first chapter of this thesis we studied the impact of civil war on human capital as captured by the anthropometry of children less than 5 years, school enrolment among children at primary school, expenditure per adult equivalent, and fertility. Before empirically testing the impact of the Angolan civil war, we constructed a theoretical model showing how civil war may affect human capital. While the choice of the location of conflict was not random, we used an instrumental variables method to assess the impact of the civil war on the human capital, expenditures per adult equivalent and fertility in Angola. Our results show that in the short term, civil war has a disastrous impact on the anthropometry of children under 5 years. However school enrolment is higher in communities most affected by civil war. The war has no short term effect on expenditure per adult equivalent and fertility. In the long term, we find a persistent and negative effect on children under 5 years. In the long term, we find no impact of civil war on education, expenditure per adult equivalent, and fertility in Angola. This chapter contributes to the new microeconomic literature of civil war on the human capital. In this chapter we go beyond the empirical estimation by constructing a theoretical model. The construction of a theoretical model also helps identify the channels through which war affects human capital and may also be useful in formulating appropriate post-conflict strategies. Our results suggest the usefulness of establishing, during or after conflict, health and nutrition programs to mitigate the effects of war on children under 5 years. The lack of impacts on education and expenditure per equivalent adult does not mean that the war had no effect on enrolment and expenditures per adult equivalent in Angola. This result simply means that there is no distinction in the long term in terms of education among communities more affected by war and communities less affected by war. We cannot evaluate the impact at the aggregate level of war in Angola because we do not have a counterfactual, but other studies show that countries in conflict or post-conflict countries have an average life expectancy 12 years lower and an enrollment rate much lower than the least developed countries. All this leads us to say that it is likely that the rate of enrollment or expenditure per adult equivalent in Angola would have been greater in the absence of civil war. Thus, despite the lack of a long-term impact of on education and expenditure per adult equivalent, it may still be important in the context of post-conflict strategies to increase access to education and to establish programs capable of generating income.

In Chapter two of this thesis, we studied the impact of the Angolan social fund on the anthropometry of children under 5 years and the expenditure per adult equivalent. We have derived our identification strategy based on a model of political economy. Our results show that the Angola social fund has a strong effect on both measures. We also show that the program's impact on expenditures per adult equivalent is higher in communities that have been most affected by war. The impact of the Angolan social fund on our outcomes is achieved primarily through the observed increases in income for households generated by this program. This result shows that despite many challenges facing countries in conflict and post-conflict, programs such as social funds can be used both to mitigate the impact of civil war on the human capital and to generate additional income for households as a source of lasting stability. This result also shows the value of targeting in conflict or post-conflict to maximize the benefit of the program. The results, in the absence of program heterogeneity among different subgroups, suggest that targeting based solely on the intensity of the conflict in a community helps to mitigate the impact of civil war, which means that before the establishment of a social fund-type program in a country in conflict or post-conflict, it is useful to design a map of conflict intensity. An important part of the impact of the Angolan social fund is growing revenue, implying that the design of the program allows households to restore their livelihoods during or after conflict.

The question that remains to be addressed is on the sustainability of such programs, as they are mostly funded by donors seeking to maximize impact. This underlines the importance of demonistrating the staying power of income-generating agricultural activities that could allow households to support themselves over time. This raises the issue of land securitization and the return of displaced persons, as all of these aspects must be integrated into the global context of post-conflict assistance in order to mitigate the negative impact of a country's civil war but also to lay the foundations for its sustainable development. It is important to have a global approach in the context of countries in confict or post-confict periods.

In the last chapter of this thesis we studied the impact of teacher training on HIV/AIDS on the level of knowledge, attitudes, and behaviors among young students in secondary schools in Cameroon. While the assignment of the program was not random, we use regression discontinuity design to estimate the causal effect of this program. Our results shows that 15 to 17 year old girls in treatment schools teacher training schools are between 7 and 10 percentage points less likely to have started childbearing, an objective proxy for the incidence of unprotected sex. They are also significantly more likely to have used a condom during their last sexual intercourse. For 12 to 13 year old girls, the likelihood of self-reported abstinence and condom use is also significantly higher in treated schools, while the likelihood of having multiple partners is significantly lower. We showed that this result can be explained by the fact that teachers have focused on reducing risk behavior during sex, defense options, and discussions around condom use. This demonstrates the importance of teaching content in HIV/AIDS awareness campaigns in schools. Despite widespread teachings on HIV/AIDS in sub-Saharan Africa, the courses' content are quite diverse, ranging from situations in which, for instance, discussion about condoms is allowed to situations were such discussion is prohibited under the assumption that merely discussing contraception will encourage early sex. All this assumes that it is important to have a certain uniformity in teaching programs related to HIV/AIDS in sub-Saharan Africa to ensure the effectiveness of this type of education program. The choice of lessons should be based on their effectiveness as estimated through measurable indicators, including those not dependent on self-reporting.

In the future, we intend to explore the research questions surrounding the intersection of HIV/AIDS and civil war in sub-Saharan Africa. Civil war drives the movement of populations and often produces situations of high infection risk, maybe most significantly the sexual abuse of women. Exploring the causal mechanisms linking conflict to HIV spread could illuminate the intersectional effects each phenomenon exerts upon human capital development in sub-Saharan Africa.

Bibliography

AGADJANIAN, V., AND N. PRATA (2002): "War, Peace, and Fertility in Angola," *Demography*, 39(2), 215–231.

AHUJA, A., B. WENDELL, AND E. WERKER (2006): "Male Circumcision and AIDS: The Macroeconomic Impact of a Health Crisis," Harvard School of Business, Working Paper No. 07-025, Cambridge, MA.

AKRESH, R., AND D. DE WALQUE (2008): "Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide," World Bank Policy Research Working Paper No. 4606, April.

ALDERMAN, H., J. HODDINOTT, AND B. KINSEY (2006): "Long Term Consequences of Early Childhood Malnutrition," *Oxford Economic Papers*, 58(3), 450–474.

ANGOLA PEACE MONITOR (2000): "Jamba Falls to FAA Following Heavy Fighting in Border Region," *Angola Peace Monitor*, VI(5), 18 January 2000.

ANGRIST, J. D., AND V. LAVY (1999): "Using Maimonides Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, 114(2), 533–575.

APPLETON, S. (2001): "Education, Incomes and Poverty in Uganda in the 1990s," CREDIT Research Paper No. 01/22.

Barron, P., K. Kaiser, and M. Pradhan (2004): "Local Conflict in Indonesia: Measuring Incidence and Identifying Patterns," World Bank Policy Research Working Paper, 3384.

Bell, C., R. Bruhns, and H. Gersbach (2006): "Economic growth, education, and AIDS in Kenya : a long-run analysis," Policy Research Working Paper No. 4025, World Bank, Washington, DC.

Bellows, J., and E. Miguel (2006): "War and Institutions: New Evidence from Sierra Leone," *American Economic Review, Papers and Proceedings*, 96(2), 394–399.

——— (2008): "War and Local Collective Action in Sierra Leone," Department of Economics, University of California, Berkeley, January.

Benhabib, J., and A. Rustichini (1991): "Vintage Capital, Investment, and Growth," *Journal of Economic Theory*, 55(2), 323–339.

Besley, T., and S. Coate (1992): "Workfare vs. Welfare: Incentive Arguments for Work Requirements in Poverty Alleviation Programs," *American Economic Review*, 82(1), 249–261.

Bethke, L., and S. Braunschweig (2003): *Global Survey on Education in Emergencies. Angola Country Report*. Women's Commission for Refugee Women and Children, New York, NY.

Biddlecom, A., L. Hessburg, S. Singh, A. Bankole, and L. Darabi (2007): "Protecting the next generation in sub-Saharan Africa: learning from adolescents to prevent HIV and unintended pregnancy," Guttmacher Institute.

Blattman, C., and E. Miguel (2008): "Civil War," *Journal of Economic Literature*, forthcoming, June.

BLOCK, S., AND P. WEBB (2003): "Nutrition Knowledge and Parental Schooling as Inputs to Child Nutrition in the Long and Short Run," Friedman School of Nutrition Science and Policy, Nutrition Working Paper No.3.

BLOOM, C., AND A. MAHAL (1997): "Does the AIDS Epidemic Really Threaten Economic Growth?," *Journal of Econometrics*, 77(1), 107Ű129.

BLUNDELL, R., L. DEARDEN, AND B. SIANESI (2004): "Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS," CEE Discussion Papers 0047, Centre for the Economics of Education, LSE.

BODEA, C., AND I. ELBADAWI (2008): "Political Violence and Economic Growth," World Bank Policy Research Working Paper No. 4692.

BONNEL, R. (2000): "HIV/AIDS: Does it increase or decrease growth in Africa?," AIDS Campaign Team. World Bank ,Washington, D.C.

BOULDING, K. (1962): *Conflict and Defense*. Harper and Row, New York, NY.

BRAKMAN, S., H. GARRETSEN, AND M. SCHRAMM (2004): "The Strategic Bombing of Cities in Germany in World War II and It Impact on City Growth," *Journal of Economic Geography*, 4(1), 1–18.

BUHAUG, H. (2007): "Dude, Where's My Conflict? LSG, Relative Strength, and the Location of Civil War," Centre for the Study of Civil War (CSCW) and International Peace Research Institute, Oslo (PRIO), February.

BUHAUG, H., AND P. I. LUJALA (2005): "Accounting for Scale: Measuring Geography in Quantitative Studies of Civil War," *Political Geography*, 24(4), 399–418.

BUHAUG, H., AND J. K. RØD (2006): "Local Determinants of African Civil Wars, 1970-2001," *Political Geography*, 25(3), 315–335.

Bundervoet, T., R. Akresh, and P. Verwimp (2008): "Health and Civil War in Burundi," *Journal of Human Resources*, forthcoming.

Cameron, A. C., G. Jonah B., and D. L. Miller (2006): "Robust Inference with Multi-way Clustering," NBER Technical Working Paper No. 327.

Carneiro, P., J. J. Heckman, and E. Vytlacil (2003): "Understanding What Instrumental Variables Estimate: Estimating Marginal and Average Returns to Education," processed, University of Chicago, The American Bar Foundation and Stanford University, July 19.

Case, A., and A. Carly (2006): "The Impact of Parental Death on School Outcomes: Longitudinal Evidence From South Africa," *Demography*, 43(3), 401–420.

Chase, R. S. (2002): "Supporting Communities in Transition: The Impact of the Armenian Social Investment Fund," *World Bank Economic Review*, 16(2), 219–240.

Chase, R. S., and L. Sherbume-Benz (2001): "Household Effects of Community Education and Health Initiatives: Evaluating the Impact of the Zambia Social Fund," The World Bank, processed, January.

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

Collier, P. (1999): "On the Economic Consequences of Civil War," *Oxford Economic Papers*, 51(1), 168–183.

——— (2007): *The Bottom Billion*. Oxford University Press, Oxford, 1th edn.

Collier, P., and A. Hoeffler (1998): "On Economic Causes of Civil War," *Oxford Economic Papers*, 50(4), 563–573.

——— (2004): "Greed and Grievance in Civil War," *Oxford Economic Papers*, 56(4), 563–595.

Collier, P., and N. Sambanis (eds.) (2005): *Understanding Civil War: Evidence and Analysis*, vol. 1: Africa. The World Bank, Washington, D.C.

Corrigan, P., G. Glomm, and F. Mendez (2005): "AIDS Crisis and Growth," *Journal of Development Economics*, 77(1), 107–131.

David, V., M. Moncada, and F. Ordonez (2004): "Private and public determinants of child nutrition in Nicaragua and Western Honduras," *Economics and Human Biology*, 2(3), 457Ű488.

Davis, D. R., and D. Weinstein (2002): "Bones, Bombs, and Breakpoints: The Geography of Economic Activity," *American Economic Review*, 92(5), 1269–1289.

De Walque, D. (2006): "Who Gets AIDS and How? The Determinants of HIV Infection and Sexual Behaviors in Burkina Faso, Cameroon, Ghana, Kenya, and Tanzania," World Bank Policy Research Working Paper No. 3844.

Dias Amaral, J. G. (2004): "Angola: A Crise Económica Na Primeira República — 1975-1992," Luanda, Centro de Estudos da Universidade Católica de Angola.

Dilolwa, C. R. (1978): *Contribuição À História de Angola*. Imprensa Nacional de Angola, Luanda, Angola.

Duflo, E., P. Dupas, M. Kremer, and S. Sinei (2006): "Education and HIV/AIDS Prevention: Evidence from a Randomized Evaluation in Western Kenya," World Bank Policy Research Working Paper No. 4024.

Dupas, P. (2009): "Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya," NBER Working Papers 14707.

EBOKO, F. (2006): "L'organisation de la lutte contre le sida au Cameroun : de la verticalité a la dispersion?," Un Systeme de Sante En Mutation: Le Cas Du Cameroun, Le bulletin de APAD no. 21.

ECONOMIDES, N. (1986): "Nash Equilibrium in Duopoly with Products Defined by Two Characteristics," *Rand Journal of Economics*, 17(3), 431–439.

EVIA, J. . L., J. NEWMAN, M. PRADHAN, C. RAMIRO, L. B. RAWLING, AND G. RIDDER (2002): "An Impact Evaluation of Education, Health, and Water Supply Investments by the Bolivian Social Investment Fund," *World Bank Economic Review*, 16(2), 241–274.

FEARON, J. D., AND D. LAITIN (2003): "Ethnicity, Insurgency and Civil War," *American Political Science Review*, 97(1), 75–90.

FROLICH, M., AND R. VAZQUEZ-ALVAREZ (2009): "HIV/AIDS Knowledge and Behaviour: Have Information Campaigns Reduced HIV Infection? The Case of Kenya," *African Development Review*, 21(1), 86–146.

GALIANI, S., P. GERTLER, AND E. SCHARGRODSKY (2005): "Water for Life: The Impact of Privatization of Water Services on Child Mortality," *Journal of Political Economy*, 113(1), 83–120.

GALLANTA, M., AND E. MATICKA-TYNDALE (2004): "School-Based HIV Prevention Programmes for African Youth," *Social Science and Medicine*, 58, 1337–1351.

GARCIA MÁRQUEZ, G. (1985): *El Amor En Los Tiempos Del Cólera*. Oveja Negra, Bogotá, Colombia.

GERSHENSON, D., AND H. I. GROSSMAN (2000): "Civil Conflict : Ended or Never Ending?," *Journal of Conflict Resolution*, 44(6), 808–822.

Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002): "Armed Conflict 1946-2001: A New Dataset," *Journal of Peace Research*, 39(5), 615–637.

GOA (2005): "Estudo de Caso de Angola Sobre O Direito Humano À Alimentação Adequada," Luanda, Angola.

Grossman, H. I. (1991): "A General Equilibrium Model of Insurrections," *American Economic Review*, 81(4), 912–921.

——— (1992): "Foreign Aid and Insurrection," *Defence and Peace Economics*, 3(4), 275–288.

——— (1999): "Kleptocracy and Revolutions," *Oxford Economic Papers*, 51(2), 267–283.

Guerrero-Serdan, G. (2009): "The Effects of the War in Iraq on Nutrition and Health:An Analysis Using Anthropometric Outcomes of Children," Households in Conflict Network Working Paper No. 55, February.

Guidolin, M., and E. La Ferrara (2007): "Diamonds Are Forever, Wars Are Not: Is Conflict Bad for Private Firms?," *American Economic Review*, 97(5), 1978–1993.

Haacker, M. (2004): "HIV/AIDS: The Impact on the Social Fabric and the Economy," in *In The Macroeconomics of HIV/AIDS*, ed. by M. Haacker, pp. 198–258, Washington, DC. International Monetary Fund.

Hagelstein, R. (2008): "Explaining the Violence Pattern of the Algerian Civil War," Households in Conflict Network Working Paper No. 43, March.

Hahn, J., P. Todd, and W. Van Der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201–209.

Hausman, J. A., and W. Taylor (1981): "Panel Data and Unobservable Individual Effects," *Econometrica*, 49(6), 1377–1398.

Hayes, R., J. Changaluchab, D. Ross, A. Gavyole, J. Todd, A. Obasi, M. Plummer, D. Wight, D. Mabey, and H. Grosskurth (2005): "The MEMA Kwa Vijana Project: design of a community randomised trial of an innovative adolescent sexual health intervention in rural Tanzania," *Contemporary Clinical Trials*, 26(4), 430–442.

Heckman, J. J. (2001): "Micro Data, Heterogeneity and the Evaluation of Public Policy: Nobel Lecture," *Journal of Political Economy*, 109(4), 673–748.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006): "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, 88(3), 389–432.

Heckman, J. J., and E. J. Vytlacil (1999): "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects," *Proceedings of the National Academy of Sciences*, 96(8), 4730–4734.

——— (2005): "Structural Equations, Treatment Effects and Econometric Policy Evaluation," *Econometrica*, 73(3), 669–738.

Hegre, H., G. Ostby, and C. Raleigh (2007): "Economic Deprivation and Civil War Events: Disaggregated Study of Liberia," Paper presented at the 48th Annual Convention of the International Studies Association, Chicago.

Hegre, H., and C. Raleigh (2007): "Population Size, Concentration, and Civil War. A Geographically Disaggregated Analysis," World Bank Policy Research Working Paper No. 4243 (Post-Conflict Transitions Working Paper No. 16).

Hobcraft, J. (1996): "Fertility in England and Wales: A Fifty Year Perspective," *Population Studies*, 50(3), 485–524.

Hoddinott, J., and B. Kinsey (2001): "Child Growth in the Time of Drought," *Oxford Bulletin of Economics & Statistics*, 63(4), 409–445.

Hotelling, H. (1929): "Stability in Competition," *Economic Journal*, 39(153), 41–57.

Humphreys, M., and J. Weinstein (2008): "Who Fights? The Determinants of Participation in Civil War," *American Political Science Review*, 52(2), 436–455.

Imbens, G. W., and T. Lemieux (2008): "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142(2), 615–635.

Jarrow, R., D. Ruppert, and Y. Yu (2004): "Estimating the Term Structure of Corporate Debt with a Semiparametric Penalized Spline Model," *Journal of the American Statistical Association*, 99(465), 57–66.

Justino, P. (2007): "On the Links Between Violent Conflict and Household Poverty: How Much Do We Really Know?," MICROCON Research Working Paper No. 1.

Kalemli-Ozcan, S. (2006): "AIDS, Reversal of the Demographic Transition and Economic Development: Evidence for Africa," NBER Working Paper No. 12181, National Bureau of Economic Research, Cambridge, MA.

Khlat, M., M. Deeb, and Y. Courbage (1997): "Fertility Levels and Differentials in Beirut During Wartime: An Indirect Estimation Based on Maternity Registers," *Population Studies*, 51(1), 85–92.

Kirby, D., B. Laris, and L. Rolleri (2007): "Sex and HIV education programs: their impact on sexual behaviors of young people throughout the world," *Journal of Adolescent Health*, 40(3), 206–217.

Kulczycki, A., and P. C. Saxena (2000): "New Evidence on Fertility Transition Through Wartime in Lebanon," paper presented at the Population Association of America Annual Meeting, Los Angeles, California, 23-25 March, 2000.

LADIER-FOULADI, M., AND H. HOURCADE (1997): "The Paradoxes of Demographic Trends in Iran: Fertility in Rapid Decline," in *La Population Du Monde: Enjeux et Problèmes*, ed. by J.-C. Chasteland, and J.-C. Chesnais, pp. 209–224, Paris. Presses Universitaires de France.

LEE, D., AND D. CARD (2008): "Regression discontinuity inference with specification error," *Journal of Econometrics*, 142(2), 655–674.

LEE, D. S., AND T. LEMIEUX (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(2), 281–355.

LEIPZIGER, D., M. FAY, Q. WODON, AND T. YEPES (2005): "Achieving Child Health Related Millenium Development Goals: The Role of Infrastructure," *World Development*, 33(8), 1267–1284.

LEVY, D., AND J. OHLS (2007): "Evaluation of Jamaica's PATH Programme: Final Report," Mathematica Policy Research, Inc. MPR Reference No. 8966-090, March 2007.

LINDSTROM, D. P., AND B. BERHANU (1997): "The Impact of War, Famine, and Economic Decline on Marital Fertility in Ethiopia," *Demography*, 36(2), 247–261.

LIPTON, M.AND M., R. (1995): "Poverty and Policy," in *Handbook of Development Economics Volume*, ed. by J. Behrman, and S. T.N., vol. 3, Amsterdam: North-Holland.

LORENTZEN, P., J. MCMILLAN, AND R. WACZIARG (2008): "Death and development," *Journal of Economic Growth*, 13(2), 81Ű124.

MACQUEEN, N. (1997): *The Decolonization of Portuguese Africa*. Longman, London, England and New York, NY.

MANSURI, G., AND R. VIJAYENDRA (2004): "Community-Based and -Driven Development: A Critical Review," *World Bank Research Observer*, 19(1), 1–39.

Marques, W. (1963): *Problemas Do Desenvolvimento Económico de Angola.* Junta de Desenvolvimento Industrial, Luanda, Angola.

Maticka-Tyndale, E., and C. Brouillard-Coyle (2006): "The effectiveness of community interventions targeting HIV and AIDS prevention at young people in developing countries," in *Preventing HIV/AIDS in young people: a systematic review of evidence from developing countries*, ed. by D. Ross, B. Dick, and J. Ferguson. UNAIDS Inter-Agency Task Team on Young People. Geneva, World Health Organization.

Miguel, E., and G. Roland (2005): "The Long Run Impact of Bombing Vietnam," Department of Economics, University of California, Berkeley, November.

Miguel, E., S. Satyanath, and E. Sergenti (2004): "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy*, 112(4), 725–753.

MINESEC (2006): "Curriculum de formation et enseignement: Education à la Vie Familiale, En Matiere de Population et Au VIH/SIDA (EVF/EMP/VIH et Sida)," Minesec, Yaounde, Cameroun.

Murshed, S. M., and S. Gates (2005): "Spatial-Horizontal Inequality and the Maoist Insurgency in Nepal," *Review of Development Economics*, 9(1), 121–134.

National Research Council, I. o. M. (2005): "Growing Up Global: The Changing Transitions to Adulthood in Developing Countries," National Academies Press.

Newell, J., and J. Einbeck (2007): "A Comparative Study of Nonparametric Derivative Estimators," Proceedings of the 22th IWSM, Barcelona.

Newey, W. K., and J. L. Powell (2003): "Instrumental Variable Estimation of Nonparametric Models," *Econometrica*, 71(5), 1565–1578.

Oster, E. (2009): "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa," NBER Working Paper Number 13610.

Over, M. (1998): "The effects of societal variables on urban rates of HIV infection in developing countries: An exploratory analysis," in *Confronting AIDS: Evidence from the Developing World*, ed. by M. Ainsworth, and M. Fransen, L.and Over, pp. 39–51, Luxembourg. European Communities.

Owen, D., and I. Bannon (2006): "Community-Driven Development in the Context of Conflict-Affected Countries: Challenges and Opportunities," World Bank, Washington D.C., June.

Papageorgiou, C., and P. Stoytcheva (2005): "What Do We Know About the Impact of AIDS on Cross-Country Income So Far?," Departmental Working Papers, Department of Economics, Louisiana State University, January.

Paxson, C., and N. R. Schady (2002): "The Allocation and Impact of Social Funds: Spending on School Infrastructure in Peru," *World Bank Economic Review*, 16(2), 297–319.

Pradhan, and L. B. Rawlings (2002): "The Impact and Targeting of Social Infrastructure Investments: Lessons from Nicaraguan Social Fund," *World Bank Economic Review*, 16(2), 275–295.

Ramos, L., and C. Siamatowe (2006): "HIV and AIDS Education: Teacher Training and Teaching," Paris, France: UNESCO.

Ravallion, M. (1991): "Reaching the Rural Poor Through Public Employment: Arguments, Evidence and Lessons from South Asia," *World Bank Research Observer*, 6(2), 153–175.

——— (1999): "Appraising Workfare," *World Bank Research Observer*, 14(1), 31–48.

Rindfuss, R. R., and J. A. Sweet (1977): *Postwar Fertility Trends and Differentials in the United States*. Academic Press, New York, NY.

Robinson, W. S. (1950): "Ecological Correlations and the Behavior of Individuals," *American Sociological Review*, 15(X), 351–357.

Rodrigues, C. U. (2003): "Recomposição Social e Urbanização Em Luanda," *Centro de Estudos Africanos Occasional Papers*, Lisbon, Portugal.

Rodrik, D. (1999): "Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses," *Journal of Economic Growth*, 4(4), 385–412.

Rosenblum, M., N. Jewell, M. Van Der Laan, S. Shiboski, A. Van Der Straten, and N. Padian (2009): "Analysing direct effects in randomized trials with secondary interventions: an application to human immunodeficiency virus prevention trials," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(2), 443–465.

Rosenzweig, M. R. (1990): "Population Growth and Human Capital Investments: Theory and Evidence," *Journal of Political Economy*, 98(5), S38–S70.

Ruppert, D., M. P. Wand, and R. J. Carroll (2000): *Semiparametric Regression*. Cambridge University Press, New York, NY.

Schwartz, K. (1997): "One Hundred Years of Fertility Development," *Zeitschrift fur Bevolkerungswissenschaft*, 22(4), 481–491.

Shemyakina, O. (2006): "The Effect of Armed Conflict on Accumulation of Schooling: Results from Tajikistan," Households in Conflict Network Working Paper No. 12.

Spears, I. A. (1999): "Angola's Elusive Peace: The Collapse of the Lusaka Accord," *International Journal*, 54(4), 562–581.

Stanton, B., X. Li, J. Kahihuata, A. Fitzgerald, S. Neumbo, G. Kanduuombe, I. Ricardo, J. Galbraith, N. Terreri, I. Guevara, H. Shipena, J. Strijdom, R. Clemens, and R. Zimba (2009): "Increased protected sex and abstinence among Namibian youth following a HIV risk-reduction intervention: a randomized, longitudinal study," *AIDS*, 12(18), 2473–2480.

Stillwaggon, E. (2000): "HIV Transmission in Latin America: Comparisons with Africa and Policy Implications," *South African Journal of Economics*, 68(5), 444–454.

Thomas, D., and J. Strauss (1997): "Health and Wages: Evidence on Men and Women in Urban Brazil," *Journal of Econometrics*, 77(1), 159–185.

Thornton, R. L. (2008): "The Demand for, and Impact of, Learning HIV Status," *American Economic Review*, 98(5), 1829–1863.

Tvedten, I. (2002): "Angola 2001/2002 Key Development Issues and Aid in a Context of Peace," *Report R 2002:8*, Chr. Michelsen Institute Development Studies and Human Rights, Oslo, Norway.

UN (2002): *Angola: The Post-War Challenges. Common Country Assessment 2002.* United Nations System in Angola.

UNAIDS (2008): "Report on the global HIV/AIDS epidemic," The Joint United Nations Programme on HIV/AIDS, Geneva, Switzerland.

UNICEF (2003): *MICS Multiple Indicator Cluster Survey. Assessing the Situation of Angolan Children and Women at the Beginning of the Millennium. Analytical Report.*

USAID (2002): *Angola Transition and Development Assessment.* U.S. Agency for International Development, Washington, D.C.

Valdivia, M. (2004): "Poverty, health infrastructure and the nutrition of Peruvian Children," *Economics & Human Biology*, 2(3), 489–510.

Verwimp, P. (2005): "An economic profile of peasant perpetrators of genocide Micro-level evidence from Rwanda," *Journal of Development Economics*, 77(2), 297–323.

Wassenich, P., and K. Whiteside (2004): "CDD Impact Assessments Study: Optimizing Evaluation Design Under Constraints," SOCIAL DEVELOPMENT PAPERS, Number 51/February 2004.

WHO (2006): "Preventing HIV/AIDS in young people: a systematic review of the evidence from developing countries," WHO Technical Report Series No. 938. August. World Health Organization, Geneva.

Woldemicael, G. (2005): "Evidence of Recent Fertility Decline in Eritrea: An Analysis of Trends and Determinants," Max Planck Institute for Demographic Research WP 2005-026.

Wooldridge, J. (2002): *Econometric Analysis of Cross-Section and Panel Data*. MIT Press, Cambridge, MA, 1st edn.

Ziemke, J. (2007): "From Battles to Massacres: An Analysis of Changing Conflict Patterns in Angola: 1961-2002," University of Wisconsin-Madison.

ABSTRACT

This thesis is based on three essays. The first chapter analyses the impact of 27 years of civil war in Angola on human capital, expenditures per adult equivalent and fertility. The prediction of the effects of civil war is done through a neoclassical unitary household model in the tradition of Rosenzweig. Using instrumental variable method, this thesis shows that civil war has a negative and disastrous impact in short-term on health of children, this effect is persistent. Civil war has no impact on expenditures per adult equivalent. It increases enrollment and decreases fertility in the short term. The second chapter of this thesis analyzes the effectiveness of a social program in a conflict country such as Angola and explores whether this effectiveness depends on the intensity of the conflict. Our identification strategy is based on the political geography of the deployment of the program based on a model of spatial competition of Hotelling. This thesis shows that the Angola Social Fund had a positive impact on expenditures per adult equivalent and on one of the main anthropometric measurements namely the height for age z-score. The program's effectiveness in function to the intensity of the conflict is analyzed using the local instrumental variable estimator. The thesis shows that the program's effectiveness increases with the intensity of the conflict. The last chapter of this thesis analyzes in the case of Cameroon, the impact of teacher training on HIV/AIDS. The two criteria for selecting participating schools, leads us to choose as identification strategy the regression discontinuity design. This thesis shows that 15 to 17 year old girls in teacher training schools are between 7 and 10 percentage points less likely to have started childbearing. For 12 to 13 year old girls, the likelihood of self-reported abstinence and condom use is also significantly higher in treated schools.
Keywords: Angola, Civil war, Instrumental variable, Local instrumental variable, Anthropometric, Cameroon, HIV/AIDS, Regression discontinuity design.

RÉSUMÉ

Cette thèse s'articule autour de trois essais. Le premier chapitre analyse l'impact des 27 ans de guerre civile en Angola sur les dépenses par équivalent adulte, le capital humain et la fécondité. La prédiction des effets de la guerre se fait à l'aide d'un modèle néoclassique de ménage unitaire dans la tradition de Rosenzweig. A partir de l'approche d'estimation par variable instrumentale, cette thèse montre que la guerre civile a un impact négatif et désastreux à court terme sur la santé des enfants, cet effet est persistant. La guerre civile n'a pas d'impact sur les dépenses par équivalent adulte. Elle accroit la scolarisation et décroit la fécondité à court terme. Le second chapitre de cette thèse analyse l'efficacité d'un programme social dans un environnement en conflit comme celui de l'Angola et s'interroge si cette efficacité dépend de l'intensité du conflit. Notre stratégie d'identification est basée sur la géographie politique du déploiement du programme basée sur un modèle de compétition spatiale à la Hotelling. Cette thèse montre que le Fond Social Angolais a eu un impact positif sur les dépenses par équivalent adulte et sur l'une des principales mesures anthropométriques à savoir le z-score de la taille pour âge. L'efficacité du programme en fonction de l'intensité du conflit est analysée à l'aide de l'estimateur de variable instrumental local. La thèse montre que l'efficacité du programme augmente avec l'intensité du conflit. Le dernier chapitre de cette thèse analyse dans le cas du Cameroun, l'impact de la formation des enseignants en matière de VIH/SIDA. Les deux critères retenus pour le choix des écoles participant au programme, nous amène à choisir comme stratégie d'identification la régression discontinue. Cette thèse montre que les filles âgées de 15 à 17 ans dans les écoles traitées sont moins susceptibles d'avoir une grossesse involontaire. Pour les élèves âgés de 12 à 13 ans, la probabilité d'abstinence auto déclarée et l'utilisation du préservatif est également significativement plus élevé dans les écoles traitées.
Mots-clefs : Angola, Guerre civile, Estimation par variable instrumentale, Estimateur de variable instrumental local, Anthropométrie, Cameroun, VIH/SIDA, Modèle de Régression discontinue.

Mots-clefs : Angola, Guerre civile, Estimation par variable instrumentale, Estimateur de variable instrumental local, Anthropométrie, Cameroun, VIH/SIDA, Modèle de Régression discontinue.