Title
Essays in East African Development

Author
Friedman, Willa Helterline

Publication Date
2013

Peer reviewed|Thesis/dissertation

Essays in East African Development

By

Willa Helterline Friedman

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

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Edward Miguel, Chair

Professor Frederico Finan

Professor Ernesto Dal Bó

Spring 2013

Abstract

Essays in East African Development

by

Willa Helterline Friedman

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Edward Miguel, Chair

This dissertation combines three studies from East Africa, each investigating a different way in which individuals respond to institutional changes, and how this response shapes the broader impact of these policy levers. Each paper develops and explains one or more theories of individual behavior and combines data from a wide range of sources to test hypotheses generated from the original theories.

The first chapter measures the impact of provision of ARVs on risky sexual behavior among young women in Kenya, and estimates how this response will shape future HIV infection rates. Access to antiretroviral (ARV) drugs in Sub-Saharan Africa has rapidly expanded - from fewer than 10,000 people treated in 2000 to more than 8 million in 2011. To measure the impact of this expansion, it is necessary to identify the behavioral response of individuals to drug access. This chapter combines geocoded information about the timing of introduction of ARVs in all Kenyan health facilities with two waves of geocoded population surveys to estimate the impact of proximity to an ARV provider on risky sexual behavior. Using a difference in differences strategy that matches survey clusters geographically across waves, I find a relative increase in risky behavior as reflected in pregnancy rates (increase of 82%) and self-reported recent sexual activity (increase of 40%) among young women in areas in which ARVs were introduced between 2004 and 2008. The full impact of ARV access on new infections is estimated through a simulation procedure that combines estimated behavioral responses to ARVs with medical evidence regarding HIV transmission. An increase in ARV drug access is predicted to reduce the rate of new infections despite the induced increase in risk-taking.

The second chapter looks at the impact of corruption on the effectiveness of antiretroviral drugs in preventing deaths due to HIV and the potential channels that generate this relationship. This is based on a unique panel dataset of countries in sub-Saharan Africa, which combines information on all imported antiretroviral drugs from the World Health Organization's Global Price Reporting Mechanism with measures of corruption and estimates of the HIV prevalence and the number of deaths in each year and in each country. Countries with higher levels of corruption experience a significantly smaller drop in HIV deaths as a result of the same quantity of ARVs imported. This is followed up with a single case-study from Kenya to illustrate one potential mechanism for the observed effect, demonstrating that disproportionately more clinics begin distributing ARVs in areas that are predominantly represented by the new leader's ethnic group.

The third chapter uses new data on participation to examine how local economic condi-

tions shaped within-country variation in willingness to participate in violent activities during the Rwandan genocide. It discusses and tests the predictions of three sets of theories about the causes of violence. The data provide strong evidence that higher rates of both unemployment and education among Hutu are associated with increased participation. I find no evidence that the employment or education of the Tutsi population reduce participation rates. I also find suggestive evidence of a positive association between violence and the interaction of Hutu unemployment and education both at the commune level and at the individual level. These results are consistent with theories of opportunity costs discouraging violence, and they provide additional evidence of a connection between education, unemployment, and violence.

To
Dancille from Ntarama,
Sharon from Busitema,
Alice from Kisoro, and
Maria-fitini from Kati

Contents

1 Antiretroviral Drug Access and Behavior Change **1**
- 1.1 Introduction . . . 1
- 1.2 Theoretical Framework . . . 5
- 1.3 Data and Context . . . 10
- 1.4 Empirical Strategy . . . 12
- 1.5 Results . . . 14
- 1.6 Simulation . . . 16
- 1.7 Conclusion . . . 18
- 1.8 Figures . . . 20
- 1.9 Tables . . . 26

2 Corruption and the Effectiveness of Imported Antiretroviral Drugs in Averting HIV Deaths **33**
- 2.1 Introduction . . . 33
- 2.2 Literature review . . . 35
  - 2.2.1 Clientelism and political patronage . . . 36
  - 2.2.2 Examples from popular press . . . 37
- 2.3 Data . . . 38
  - 2.3.1 Cross-country impacts of corruption and ARVs . . . 38
  - 2.3.2 Case study . . . 39
- 2.4 Empirical Strategy and Results . . . 40
  - 2.4.1 Cross-country impacts of corruption and ARVs . . . 40
- 2.5 Case study in Kenya . . . 42
- 2.6 Conclusion . . . 44
- 2.7 Figures . . . 45
- 2.8 Tables . . . 46

3 Local Economic Conditions and Participation in the Rwandan Genocide **53**
- 3.1 Introduction . . . 53
- 3.2 Theoretical Framework and Literature Review . . . 54
  - 3.2.1 Opportunity Cost Theories . . . 55
  - 3.2.2 Expropriation and Relative Deprivation theories . . . 57
  - 3.2.3 Political Participation and Unmet Expectations . . . 59
  - 3.2.4 Literature on Violence and Civil Unrest in Rwanda . . . 60
- 3.3 Historical Context . . . 61

3.4 Data 62
3.4.1 Participation Rates 63
3.4.2 Baseline Local Economic Conditions 64
3.4.3 Combining Data 66
3.5 Empirical Strategy 66
3.6 Results and Discussion 67
3.6.1 Estimates 67
3.6.2 Discussion 68
3.6.3 Further Analysis 69
3.7 Conclusion 70
3.8 Figures 72
3.9 Tables 75

Acknowledgements

My advisors have been extremely helpful with my dissertation and my becoming an economist. In addition to providing valuable advice and encouragement, Ted Miguel has been and will continue to be an inspiration as both a researcher and a person. I was very lucky that Fred Finan arrived at Berkeley just in time to provide generous and direct feedback, which has strengthened each chapter of this dissertation enormously. Ernesto Dal Bó's enthusiasm and passion for all questions has repeatedly sent me from his office with more curiosity and energy than I had come in with.

I am also very grateful for the advice and support that I have received along the way from many other professors, including Jeremy Magruder, Justin McCrary, Dan Posner, Steve Raphael, Gerard Roland, and Ann Swidler.

These projects could not have happened without funding from IGERT; the Center for African Studies; the Institute of International Studies; and the Levin and Weiss Families and the Center for Effective Global Action. I cannot imagine a graduate advisor better than Patrick Allen.

Thank you to Jean de Dieu Bizimanga, Evan Klaus, and everyone else who helped me track down administrative records. Thank you to Dan Bennett, Damien de Walque, Jane Fortson, Erick Gong, Elaine Liu, Nicholas Wilson, Andy Zuppann, and many others for early feedback on the first chapter.

Thank you to Grace and to Megumi and Michael, and to every one else who guided and housed me in East Africa. And to anybody who ever responded to the question, "Would you mind if we come into your home and ask you a few hundred questions about your personal life for the next couple of hours?" with some variant of "Welcome to my home," I am forever indebted to you.

While we always thank seminar participants for their constructive comments, I feel that I have been particularly lucky with long-standing involvement with a few groups, each of which has provided me with incredible training, but also inspiration and friendship. This includes my IGERT fellowship-mates, six years of development lunch participants, and the incomparable Working Group on African Political Economy. WGAPE is a gem and it demonstrates what intellectual collaboration can be.

I am so grateful to many classmates and officemates and housemates and friends for what I have and will continue to learn from them, and for their kindness. Thank you to Valentina and Gabe and the rest of my first year study group for sharing sunsets in the Peixotto room and continued friendship. Thank you to Jonas and Francois and the rest of my Glen Ave housemates for breakfasts and barbecues and everything on the terrace. Thank you to Issi, Mark, and Zach and the rest of my officemates, and to Josh and Mitch and the rest of my

classmates. Thank you to Gianmarco and Josh and the rest of the development breakfast group for sharing eggs and early-stages of projects.

Thank you to Mariana, Garret, Ashley, and Owen for passing down Berkeley student's acquired wisdom, and thank you to Tadeja and Lauren for letting me pretend to pass it down and then providing amazing editing. Thank you to Ana for making the Women in Economics Group happen. It changed the whole experience. And thank you to Jamie for being amazing and for being with me in nearly every group mentioned above. I am so glad you were at Upper Hill that day in 2007.

I continue to be amazed by and grateful for the group that taught me to think and speak and listen by letting me bungle all of these as I learned: Christina, Tony, Andrew, Ishani, Matty, Buddy, and Adiari. Thank you to Lauren and Chiara for reminding me of the beauty and power of words, and thank you to Hannah, for always understanding.

Finally, thank you to my family. My father, Peter Friedman, recommended staying in school as long as possible, and my mother, Marilyn Helterline, pointed out early that economists apply math to social sciences, knowing that this would make a nice home. You were both right about these and seemingly everything else. Thank you to Sam and Eben for being great brothers. And thank you to Vikram, for patience and for love. I love you all very much.

Chapter 1

Antiretroviral Drug Access and Behavior Change

1.1 Introduction

The HIV epidemic has had an enormous impact on the well-being of millions of people in developing countries. High HIV prevalence rates are associated with falling life expectancy, substantial reductions in human capital accumulation (Cavalcanti Ferreira and Pessoa (2003), Lorentzen, McMillan and Wacziarg (2008), Fortson (2011)), reduced intergenerational human capital transmission (Beegle, De Weerdt and Dercon (2008), Bell, Devarajan and Gersbach (2004), Hunter and Williamson (2000)), and reduced economic growth (Cuddington and Hancock (1994), Corrigan, Glomm and Mendez (2005)). The introduction and rapid expansion in access to antiretroviral Drugs (ARVs), which can extend the lives of HIV positive individuals by approximately ten years, is a substantial technological innovation that has changed the course of the epidemic. While ARVs clearly benefit infected individuals and their dependents by delaying the onset of symptoms and revitalize the workforces of many developing countries, ARV provision also shapes future infection rates. Any estimation of the impact of ARVs on future HIV infections fundamentally depends on individual behavioral responses to treatment availability. The direction of this response is theoretically ambiguous because while the cost of infection has gone down, perceptions of the likelihood of infection could increase or decrease depending on beliefs about the impact of ARVs on transmission probabilities. As these beliefs cannot be observed directly, the behavioral response to ARV access must be measured empirically.

This paper uses an original dataset linking individual behavior from two waves of Demographic and Health Surveys (DHS) with a record of the roll-out of ARVs in Kenya to estimate how individual risk-taking responds to ARV access. Using a difference in differences framework with geographically identified survey clusters matched across rounds, I estimate the response to be an 82 percent increase in pregnancies and a 40 percent increase in self-reported risky sexual behavior in the previous 4 weeks. As resulting new infections cannot be empirically identified directly, this paper combines these estimates of the behavioral response to ARV access with medical evidence about a reduction in transmission probabilities for those taking ARVs to simulate the impact of ARV introduction on new infection rates. A suffi-

ciently high level of ARV provision can outweigh even this substantial increase in risk-taking, even with a conservative estimate of the reduction in transmission probability.

A simple theoretical framework demonstrates that the direction of the change in risk-taking in response to ARV access is ambiguous. On the one hand, models of behavioral disinhibition predict that when individuals are faced with an exogenous decrease in the riskiness associated with an activity over which they have some control, they may take on additional risk (e.g. Peltzman (1975)). In the case of ARV access, individuals who learn that treatment will be available may engage in more risky behavior. This constitutes a specific example of moral hazard associated with access to treatment, which implies that individuals with greater expected access to ARVs would be more likely to risk HIV infection than those who do not anticipate that treatment would be available. On the other hand, ARV access also changes both the true and the perceived probability of becoming infected. While ARV provision means more infected individuals are alive and presumably in the pool of potential sexual partners (Lakdawalla, Sood and Goldman (2006)), the medical literature has also demonstrated that these treated individuals have lower transmission probabilities. Perceptions of this can differ widely as some believe that there is no reduction in transmission probability, and others believe that the reduction is complete. This belief determines the direction of the change in the likelihood of becoming infected when ARVs are available.

Estimating the impact of access to ARVs on risk-taking presents a few key challenges that need to be addressed in order to obtain credible estimates. The first challenge to address is the definition of access to ARVs. Self-reported measures of awareness of ARVs introduce endogenous variation in individual characteristics. But proximity to an ARV facility provides an exogenous source of variation in treatment access.[1] Any measure of access based on proximity will inevitably incorporate some misclassification. However, proximity can be thought of as an instrument for access to information about treatment availability, although the first stage cannot be estimated with these data. I exploit detailed geographic information and use the location of respondents relative to health facilities providing ARVs as a proxy for access to treatment. Two measures of proximity are used. First, distance to the nearest facility defines proximity. The primary analysis uses a threshold of 8 kilometers (5 miles) to maximize power, although the results are robust to alternative thresholds. Alternatively, access is defined as being within the same administrative geographic division as a health facility that provides ARVs.[2]

The second challenge is to define a reasonable comparison group to serve as a counterfactual for those with access. I use a difference in differences identification strategy with geographic matching to deal with unobserved time-invariant differences across areas. As different villages were surveyed in each wave of the DHS, I use location to match observations across rounds. In the main specification, I match clusters of observations from each wave with those from the nearest clusters from the other wave. With multiple matches, this presents a reasonable counterfactual with which to estimate the treatment effect. This will be explained in more detail in Section 1.4. A simpler specification is also presented that compares within administrative divisions, using division fixed effects to address time-invariant

[1] In this context, an experiment may not be appropriate as informing people about the presence of ARVs in order to measure whether this encourages additional risk-taking would raise substantial ethical concerns.

[2] A division is the smallest administrative unit in Kenya, with an average size of 2181 km^2. The average size of divisions that are not excluded and contain at least 2 DHS clusters is 2007 km^2

unobserved differences.

The third challenge is that endogenous placement of ARV facilities raises concerns about omitted variables. Based on policy documents from the Kenyan Ministry of Health, I control explicitly for various factors that were used in targeting facilities for ARV introduction, including HIV rates, urban-rural status, and proximity to other health facilities. Difference in difference estimation addresses time-invariant differences across areas, but it relies on the assumption that in the counterfactual world without ARVs, trends in the control and treatment areas would have been comparable. I use a historical birth register to show that trends in pregnancy rates in treated and control areas were parallel for two decades before ARVs were introduced.

The fourth challenge to be addressed is that, while the outcome of interest in this study is sexual risk-taking, sexual behavior is notoriously misreported (e.g. Jamison and Karlan (2011), Minnis et al. (2009)). To address this, I rely primarily on pregnancy as a proxy for unprotected sexual activity. Pregnancy is a particularly appropriate proxy in this country for a few reasons. First, unlike in many developed countries, in Kenya, as in most of Sub-Saharan Africa, HIV is a generalized epidemic, predominantly spread through heterosexual sex. Second, while abortion exists, it is illegal, and therefore relatively less common. Indeed, the use of pregnancy as a marker of unprotected sex is a commonly used strategy (e.g. Duflo, Dupas and Kremer (2011), Dupas (2011)). Still, I will also report impacts on self-reported recent sexual activity.

I estimate a statistically significant relative increase in pregnancy and self-reported sexual activity in areas where ARVs were introduced among women aged 15-18. The point estimate of the treatment effect on pregnancies is 6.56 percentage points, or an increase of approximately 82 percent relative to the fertility rate in control areas. The impact on sexual activity in the previous four weeks is estimated to be a 5.72 percentage point increase or a change of 40 percent. I focus predominantly on this young demographic as they are the least likely to be in stable relationships, and therefore the most likely to change their willingness to have unprotected sex in response to changes in the threat of HIV infection. This result is consistent across different age and distance thresholds.

A final concern to address is the extent to which alternative mechanisms could explain the observed relationship between ARV access and fertility. In particular, a change in fertility preferences from an increase in life expectancy could generate the observed changes in pregnancy. However, if this were the case, we would expect to also see changes in fertility among married women, yet there is no evidence of a change in behavior among those who are married and no changes in other measures of fertility preferences or access to family planning. Another alternative mechanism relies on an increase in HIV testing that facilitates sero-sorting, namely matching among individuals with the same HIV infection status. I show that the results hold for those who have not been tested and that the size of the population that could sero-sort is sufficiently small that this cannot drive the primary empirical results.

It is not currently possible to empirically estimate the impact of ARV provision on new HIV infections with a purely quasi-experimental approach for reasons related to the biology of HIV transmission and infection. First, the full change in new infections will not be realized immediately, and it is therefore too soon after the introduction of ARVs to measure the full impact. Second, estimating the impact on *new* infections would require distinguishing between new and old infections. As ARVs keep those with HIV alive longer, there will

be a mechanical relationship between their introduction and the prevalence of HIV in the population, even if there is no impact on *new* infections. Yet distinguishing between new and old infections is infeasible.

A simulation, incorporating both medical evidence and the behavioral estimates of this paper can provide a reliable prediction of the impact of ARVs on new infection rates. It combines a range of estimates of the reduction in transmission probabilities found in the medical literature with a substantial increase in risk-taking as drugs are made available.[3] I find that even a conservative estimate of the reduction in transmission probabilities can outweigh the effects of a large increase in risk-taking if a sufficient fraction of those who are positive are treated, predicting reductions in HIV infection from an expansion in ARV access.

This paper provides a test of the theory of risk homeostasis (Peltzman (1975)), which posits that individuals may respond to a decrease in the riskiness of an action by increasing their choice of that action. This risk offset hypothesis is similar to theories of *behavioral disinhibition* due to changes in risk, to theories of *risk compensation*[4] mentioned in the public health literature, and *moral hazard* associated with treatment access. Previous empirical work has found evidence for risk homeostasis in the context of drivers' response to auto safety innovations (Winston, Maheshri and Mannering (2006)).

However, in the context of HIV risk-taking, empirical tests of theories of risk offsetting have found surprisingly little supporting evidence. For example, studies have found no expected responses in risk-taking from information about male circumcision and HIV risk (Godlonton, Munthali and Thornton (2011), Wilson, Xiong and Mattson (2012)).[5] Estimates of the behavioral response to HIV risk generally have found small or no impacts on sexual behavior (Oster (2012)) or fertility (Fortson (2011), Juhn, Kalemli-Ozcan and Turan (2008), Kalemli-Ozcan and Turan (2011)), although Young (2005) and Young (2007) do find a reduction in child-bearing associated with HIV prevalence.

A few recent papers have explored the impact of antiretroviral drugs on risk-taking with mixed results. Two studies in the US use variation in behavior among gay men before and after ARVs became available in the US, both finding an increase in risk-taking after their introduction (Mechoulan (2007), Papageorge (2012)). In Sub-Saharan Africa, where the overall HIV rate is higher, life-expectancy is lower, and access is still limited, ARVs may affect a wider range of outcomes with greater policy implications. In Malawi, Baranov, Bennett and Kohler (2012) use a method similar to this paper, but rely on measures of risk-taking largely determined before ARVs became available, and they find no impact using the entire sample. This paper focuses on recent behavior among those who could change their actions (young women) resulting in different findings. de Walque, Kazianga and Over (2012) study the impacts of beliefs about ARV effectiveness on risk-taking and find a behavioral response,

[3]Previous simulations undertook a similar exercise, but without estimates of either the reduction in transmission rates or of a change in behavior, they were somewhat inconclusive, although the authors suggested that an increase in risky behavior had a significant chance of outweighing the reduction in transmission probabilities (e.g.: Blower, Gershengorn and Grant (2000), Law et al. (2001)).

[4]This term is commonly used but should not be confused with risk compensation in the labor economics literature referring to increased wages paid to employees asked to undertake greater risks.

[5]Male circumcision is associated with a dramatic reduction in the risk of HIV infection (Auvert et al. (2005), Bailey et al. (2007), Gray et al. (2007)).

but relying on self-reported beliefs about ARVs introduces concerns about endogeneity.

Although other papers have estimated the impact of ARV access on those who are HIV positive (Bor (2012), Lakdawalla, Sood and Goldman (2006), Thirumurthy, Zivin and Goldstein (2008), Thirumurthy et al. (2012)), and the impacts on other outcomes including employment (McLaren (2012)), mortality risk-perceptions and productivity (Baranov, Bennett and Kohler (2012)), human capital investments (Baranov and Kohler (2012)), child health (Lucas and Wilson (n.d.)), and HIV testing (Wilson (2011)), this paper presents the first causally identified estimates of the impact of ARVs on risk-taking in a context with a generalized HIV epidemic, and this is the behavioral outcome that will determine the course of the epidemic.

This paper proceeds as follows. In section 1.2, I outline a theoretical framework to formalize the intuition driving the empirical estimation and to demonstrate how the empirical estimates will drive the final simulation. Section 4 describes the data and the context in which it was collected, and the empirical methods are outlined in section 5. Section 5 discusses the main results, and in section 6 I simulate the rate of new infections as a function of the level of ARV distribution, incorporating both mechanical impacts from the medical literature and the behavioral responses estimated in section 5. I conclude in section 7.

1.2 Theoretical Framework

The theoretical framework presented in this paper builds on the behavior change literature applied to responses to information about HIV. In an early model, Kremer (1996) argues that high HIV prevalence may dissuade those who are low-risk and least-likely to be infected from participating in sexual activity at all while causing those who are less cautious to take more risks because of the low probability of remaining negative. This can generate multiple equilibria at different risk levels. More recently, Gong (2011) shows that HIV testing changes behavior differentially for individuals with different priors about their own status, finding support in data from an early randomized offering of HIV testing in East Africa. Kerwin (2012) constructs a new model that rationalizes a type of fatalism based on previous risk-taking that can generate non-monotonic responses to changes in risk. This model helps to explain a pattern observed in Malawi in which individuals sufficiently overestimate the likelihood that they are currently infected, and stop taking precautions (e.g. Kaler (2003)).

In the framework developed in this paper, individuals from an infinite population of agents of size 1 choose whether or not to have unprotected sex by weighing the individual-specific benefit from unprotected sex against the expected costs of HIV infection. Access to treatment can change perceptions about both the likelihood of infection and the cost of becoming infected.[6]

The rate of new infections among those previously uninfected, I, is equal to the probability of infection conditional on engaging in unprotected sex, p, multiplied by the proportion of the uninfected population that chooses to do so, A_1. Treatment availability directly affects

[6]For simplicity, I assume that individuals who have access to ARVs know that they have access and that those who do not do not anticipate future access. This is plausible if proximity brings with it information about the existence of ARVs. I discuss the empirical implications of this assumption in section 1.3.

p by changing the pool of potential partners and their infectivity and indirectly affects both p and A_1 through a behavioral channel.

Individuals can be categorized into three types: 1) Type 1 is HIV negative, 2) Type 2 is HIV positive, without treatment, and 3) Type 3 is HIV positive, with treatment.

I make the following assumptions throughout:

- Each individual has full information about his or her own status. This assumption is included to make the model tractable and to focus on aspects which can be addressed in the empirics. The focus of the analysis is young women, who are likely to accurately perceive that it is unlikely that they are currently HIV positive. This population is old enough that infection from birth is nearly impossible, yet they are young enough that they have not or only recently began having sex. Up until that point, the likelihood of infection was approximately zero, although it could have recently changed. Throughout this section, the impact of weakening the assumption of full information about own status will be directly addressed.

- Each individual knows the distribution of other types among potential sexual partners, but does not observe the status of any particular potential partner.

Each individual chooses **whether to have unprotected sex** based on an individual-specific utility from unprotected sex (incorporating everything including social pressure and desire for children, etc.). Those who are HIV negative also consider the likelihood of becoming infected and the associated utility cost of infection.

Type 1: Formally, an uninfected individual will choose to have unprotected sex if:

$$\theta_i + (1-p) \cdot u^- + p \cdot u^+ \quad > \quad u^- \tag{1.1}$$

where u^- represents the continuation value of staying negative, u^+ represents the continuation value of being positive, and p represents the probability of infection from unprotected sex. θ_i is an individual-specific taste parameter, distributed with cdf, F_θ, which encompasses all non-HIV-related costs or benefits of unprotected sex relative to the alternative. The alternative can be abstinence or protected sex.[7] To be clear, this parameter can also be negative. Rewriting inequality 1 as $\theta_i > p \cdot (u^- - u^+)$, it follows that the proportion of the population that is negative (Type 1) that chooses to have unprotected sex can be written as:

$$A_1 \quad = \quad 1 - F_\theta(p \cdot (u^- - u^+)) \tag{1.2}$$

Note that ARV availability may change two components of the above equation. First, it reduces the relative cost of becoming infected, $u^- - u^+$, by extending the HIV positive life expectancy. This alone would lead to an increase in risk-taking among individuals of Type 1. However, ARV access can also affect p by changing the population of potential sexual partners. The direction of this effect is ambiguous.

[7] A number of papers have found evidence of a higher willingness to pay for unprotected sex among those who visit sex workers (e..g.: Gertler, Shah and Bertozzi (2005), Rao et al. (2003), and Robinson and Yeh (2011)).

If individuals do not know their HIV status, then the impact of ARV access will be dampened, but the sign will remain the same. If an individual believes that the probability he or she is HIV positive is π, then inequality 1.1 can be rewritten as

$$\theta_i + (1-\pi) \cdot [(1-p) \cdot u^- + p \cdot u^+] + \pi \cdot u^+ > (1-\pi) \cdot u^- + \pi \cdot u^+ \tag{1.3}$$

Although this changes the threshold of θ_i over which an individual chooses to have sex, it does not change the direction of the effect of ARV access via the probability of infection from sex or the cost of infection. If, however, drugs change whether people get tested for HIV, then this raises a further complication which is addressed later in the paper.

Types 2 and 3: Those who are already HIV positive do not risk changing their HIV status, and thus the only parameter in their utility optimization is the individual-specific utility from unprotected sex.[8] Altruism, morbidity, fatalism, desire for children, or any other channel through which treatment changes the utility from unprotected sex for those who are positive can be incorporated into the model by allowing this taste parameter for Types 2 and 3 to be drawn from different distributions.

Thus, an individual of Type 2 (HIV positive, not on treatment) will choose to have unprotected sex if $\gamma_i > 0$, and an individual of Type 3 (HIV positive, on treatment) will choose to have unprotected sex if $\omega_i > 0$, where γ_i and ω_i, are individual-specific taste parameters distributed with cdfs, F_γ and F_ω, respectively. These parameters can be positive or negative, incorporating any utility gains or losses from unprotected sex.

It follows that the proportion of the population that is positive and not on ARVs (Type 2) that chooses to have unprotected sex can be represented as

$$A_2 \;=\; 1 - F_\gamma(0) \tag{1.4}$$

and similarly, the proportion of the population that is positive and on ARVs (Type 3) that chooses to have unprotected sex can be represented as:

$$A_3 \;=\; 1 - F_\omega(0) \tag{1.5}$$

The assumption that those of Types 2 and 3 will not change their behavior in response to treatment access *of others* depends on the claim that while HIV positive individuals bear a utility cost from the possibility of infecting someone who is negative (i.e.: they are altruistic), altruism will have only limited behavior-change consequences. This assumption depends critically on the marginal changes in the probability that one's sexual partner is negative. Where prevalence rates in Kenya are somewhere between 5 and 15 percent, the probability of a heterogeneous match for someone who is HIV positive (i.e.: an HIV negative partner) is much higher than the probability of a heterogeneous match for someone who is HIV negative (i.e.: an HIV positive partner). Changes in the composition of the pool of potential partners induced by the medical life extension of ARVs is therefore proportionally small for those who are positive and proportionally large for those who are negative. Further,

[8] Those who are HIV positive do risk re-infection from having sex with another person who is HIV positive. This can moderately increase the speed of the progression of HIV into full-blown AIDS. However, this can credibly be assumed to be negligible with no loss to the applicability of the model.

supposing that those who are positive are very likely to draw a negative sex partner, the altruistic calculation of the cost of infecting someone on the basis of ARV availability is second order for those who are positive (who naturally discount costs for others relative to own benefits) where it is arguably quite substantial for those who are negative.

Another concern with this formulation is that many do not know their status.[9] Again, I expect this to have only a negligible effect overall. First, those of Type 3 necessarily know their status as they are receiving treatment. If some Types 1 and 2 do not know their status, then this will dampen any impact on behavior among those who are negative. As long as some of the Types 1 and 2 knows their status or the two types have different perceptions of the chance that they are infected, then the behavioral response among those of Type 2 will be smaller than those of Type 1. An increase in sexual activity among those who are positive and untreated will feedback, decreasing the utility from unprotected sex among those of Type 1. Without full information, the impact of treatment on behavior in the two types must go in the same direction. However, even if the two types share identical beliefs, this will dampen, but not change the sign of any other impacts.

The probability of becoming infected from unprotected sex, p, depends on the proportion of each type among potential sexual partners and the likelihood of transmission from each type. Denote by N_j the size of the population of each type, because the transmission probabilities can be different with these two groups.

Let q be the reduction in infectivity due to ARVs, and let $\hat{q}$ by individuals' beliefs about q.[10] If individuals believe that ARVs fully eliminate the risk of transmission, then $\hat{q}$ is 0. On the other hand, if individuals are unaware of the reduction in infectivity, then $\hat{q} = 1$.[11] For an individual of Type 1, the likelihood of infection if their partner is of Type 2 is r and the likelihood of infection if the partner is of Type 3 is $r \cdot q$.

The likelihood of infection from unprotected sex can therefore be written as:

$$p = r \cdot \frac{A_2 N_2 + A_3 N_3 q}{A_1 N_1 + A_2 N_2 + A_3 N_3} \tag{1.6}$$

and by analogy, the perceived likelihood of infection is:

$$\hat{p} = r \cdot \frac{A_2 N_2 + A_3 N_3 \hat{q}}{A_1 N_1 + A_2 N_2 + A_3 N_3} \tag{1.7}$$

Changes in access to ARVs affect p by changing the relative sizes of the population of Types 2 and 3 and the proportion of those who are negative who engage (A_1).

Let D represent the share of those who are positive who receive treatment, and let M be the share of the population that was infected as of the beginning of the current period. Besides the possibility of different behavioral parameters, γ_i and ω_i as outlined

[9] Of those who tested positive in the 2008/2009 wave of the DHS in Kenya 29 percent had never been tested for HIV previously, and so likely did not know their status.

[10] Based on the medical literature, q could be as small as 0.04 ((Cohen et al. (2011)) so the reduction in infectivity from treatment could be quite large. However, individuals respond to their beliefs, $\hat{q}$, which could be anywhere between 0 and 1.

[11] In informal conversations with HIV clinic employees, this was a commonly held belief. Many expressed concern that people who were HIV positive had become healthy and fat and were at risk of infecting others.

above, individuals of type 2 and type 3 have different death rates (d_2 and d_3 respectively), as the primary function of ARVs is to keep HIV positive individuals alive. Therefore the size of each population is:

$$\begin{aligned} N_1 &\quad \text{is fixed from the previous period.} \\ N_2 &= M \cdot (1 - D) \cdot (1 - d_2) \quad (1.8) \\ N_3 &= M \cdot D \cdot (1 - d_3) \quad (1.9) \end{aligned}$$

and we know that $d_2 > d_3$. If treatment is unavailable then $D = 0$ and $N_3 = 0$, and if everybody who is positive receives treatment, then $N_2 = 0$.

An increase in D decreases the cost of becoming infected ($u^- - u^+$), and it changes $\hat{p}$, the perceived likelihood of becoming infected. The sign of this is ambiguous and depends on other parameters.

In particular, if $\hat{q} = 0$, then:

$$\frac{d\hat{p}}{dD} < 0 \quad (1.10)$$

This is intuitive because every impact of drug provision on p moves it in the same direction. First, with the elimination of infection of those on treatment, the size of the infectious population is necessarily smaller, reducing the likelihood of matching with someone who is infectious. Second, if individuals respond to the reduction in risk from fewer positive matches or from the reduction in the cost of infection, then $A1$ will increase as well, which will further reduce p.

On the other hand, if $\hat{q} = 1$, then the impact of drugs on the likelihood of infection is more complicated. With no reduction in transmission probabilities but a reduction in the mortality probability of those infected, there will be an increase in the size of the infectious population in the pool of potential partners. This will increase p. On the other hand, if the reduction in the cost of infection sufficiently increases A_1 (the fraction of the negatives who choose to have sex), then this could reduce p. Which effect will dominate cannot be determined theoretically because it depends on the response to the perceived cost of infection. If the first effect dominates and p increases, then the effect of drugs on A_1 also becomes ambiguous.[12].

While ARV availability unambiguously decreases the cost to the individual of infection, the sign of the impact of ARV availability on the perceived probability of infection is ambiguous as is the relative magnitude of the cost reduction to the positive or negative change in the perceived probability of infection. Therefore the impact on the likelihood of those who are negative engaging in unprotected sex is ambiguous. The empirical section will estimate this revealed decision.

The theoretical framework was set up in part to show how drugs change new infections directly and through changes in behavior. As previously stated, the infection rate is:

$$I = A_1 \cdot p \quad (1.11)$$

[12]If individuals do not know their own HIV status, then A_2 will move in the same direction as A_1, which will reduce the magnitude of, but not change the sign of $\frac{d\hat{p}}{dD}$

All parameters that contribute to the above equation can be taken from the existing medical literature, with the exception of the behavioral response to treatment, which determines A_1, and indirectly, p. This response will be measured in the empirical analysis of this paper, and then this estimated response will be used to predict the impact of drugs on new infections.

1.3 Data and Context

Antiretroviral drugs were developed during the 1980s and became widely available in developed countries in the 1990s. Because of prohibitively high prices, they were almost completely unavailable to residents of Sub-Saharan Africa until the last decade. In the early 2000s, a number of agreements between developing countries and pharmaceutical companies reduced the prices of ARVs for governments of developing countries. Since then, the price of ARVs paid for by these governments has fallen from more than $10,000 per person per year to under $70 per person per year. With funding from governments and international organizations, ARVs are provided free of charge to eligible patients in Kenya and most other Sub-Saharan African countries.

As reported in Table 1.1, Kenya has a relatively high rate of HIV infection (6.3% in 2009), and it has seen a large and rapid expansion in access to ARVs in the last decade. In the early stages of the roll-out, the Ministry of Health and other associated government organizations outlined plans to provide geographically dispersed access through capable pre-existing facilities. Although initially only large hospitals were considered to have all the necessary staff and equipment to provide treatment, the requirements for facilities to be designated as capable have been reduced. In 2004, only 7 facilities distributed ARVs in Kenya but this increased substantially to 336 in 2008 (Figures 1.1a and 1.1b). Treatment is free for those who are HIV positive and eligible.[13]

Some locations were more likely than others to have ARVs introduced, and the empirical analysis will address these. This includes urban areas and areas with high rates of HIV. Because distribution happened through existing facilities, areas with large hospitals were more likely to distribute ARVs, while areas without nearby health facilities were less likely. The DHS data used in this paper provides the best existing estimates of regional HIV prevalence, and the Kenya Open Data Initiative provides a record of the GPS locations of all health facilities currently in Kenya. This information is included in the analysis to address potential endogeneity from location of ARV sources.

Information about ARV access comes from an original dataset constructed using administrative records obtained from meetings with government and NGO officials in Kenya. The geographic information comes from the Kenya Open Data Initiative,[14] and the timing information comes from reports provided by KEMSA, a procurement agency, and the National AIDS and STI Control Program (NASCOP) of the Ministry of Health. This combined database of health facilities that currently provide ARVs includes information for each

[13]Eligibility was initially based on assessments of whether a person was expected to be able to adhere to the medicine, and the progression of the disease. Now the primary metric for eligibility is the progression of the disease. Initially a person was eligible with a CD4 count below 200, but the WHO has increased the threshold to 350.

[14]See opendata.go.ke

facility on the year ARV distribution began and the location of the facility.

I hand matched clinic information across data sources by the name and district of each facility. The first instance in which a health facility appears in any records is used as the year in which treatment became available.[15] Table 1.1 shows the number of health facilities and the number of individuals receiving treatment in each year.

The data on individual behaviors come from two waves of geocoded Demographic and Health Surveys (DHS) from 2003 and 2008/2009, [16] which will be referred to throughout the paper as Wave 1 and Wave 2 respectively. Kenya expanded treatment availability largely between 2006 and 2009, so these waves provide information from before and during the middle stages of the expansion. Columns 3 and 4 of Table 1.1 shows the number of women and the number of clusters in each survey. Each cluster contains an average of 18 households and 21 female respondents. The analysis will focus on women ages 15-18 in order to look at a population that is most likely not to be in stable partnerships. Those who are already married are less likely to change their behavior in measurable ways.[17] I also exclude Nairobi and other areas which were reported to have ARV access in 2004 to mitigate concerns regarding the endogeneity of ARV access. Summary statistics of relevant variables are reported for the sample used in the analysis in Table 1.2. For clarity, all percentages are reported out of 100.

A few characteristics of the sample should be noted. First, a relatively small fraction of the sample of young women is HIV positive, but treated areas have higher prevalence rates, which will be addressed in the analysis. Nearly the entire sample in both rounds (between 97% and 100%) in both treatment and comparison areas have heard of HIV, and approximately two thirds report that they know someone who currently has or has died of AIDS. Testing increased between rounds in both areas, with a somewhat larger increase in treatment areas, which is consistent with the findings of Wilson (2011). Among both groups, only a very small fraction report STD symptoms or multiple partnerships.

The DHS data contain responses to questions about childbearing and recent sexual activity. There is extensive evidence of misreporting of sexual activity from direct survey questions (e.g. Jamison and Karlan (2011), Minnis et al. (2009)). In this particular dataset, for example, 609 women reported that the age they first had sex was later than the age at which they first gave birth, and of 2096 individuals in both waves who reported that they had never had sex, 24 tested positive for HIV. All individuals in the sample are over age 15 and therefore very unlikely to have been born with HIV, and this rate is well above the error rate of the set of tests used. Because of these concerns about measurement error, childbearing is a commonly used measure of HIV risk-taking (e.g. Duflo, Dupas and Kremer (2011), Dupas (2011)). I follow this convention and use current pregnancy as a preferred proxy for unprotected sex and show additional results using self-reported behavior as the outcome variable.[18] Results are also presented with self-reported unprotected sex in the last

[15] In conversations with officials working on Monitoring and Evaluation of ARV distribution, I was not told of any health facilities that stopped distributing drugs unless they were replaced by another organization in the same location.

[16] Interviews in the second wave were conducted between November 2008 and March 2009.

[17] While those in stable relationships may change their behavior outside of marriage in response to changes in HIV risk, this is more difficult to measure. I cannot determine paternity from the data, and only a small fraction of respondents report having additional partners. Respondents are asked about STIs, but very few report infections or symptoms.

[18] Those who report having miscarried recently and would have been pregnant (based on the number of

four weeks as the outcome.

Throughout the analysis, I proxy for information about access to ARVs with the proximity to a facility providing ARVs. A small fraction of the HIV negative population is aware of ARVs before they are introduced in the area. For this group, proximity only marginally increases access by reducing the cost of obtaining treatment. The bulk of the population has no previous information about ARVs until they are first introduced at a nearby clinic. This change in awareness can happen through several channels, including deliberate information campaigns, posters, and billboards announcing the availability of treatment.[19]

1.4 Empirical Strategy

With two waves of population surveys combined with a record of the roll-out of treatment, the estimation will rely on a difference in differences identification strategy, using multiple definitions of access based on proximity to an ARV facility and methods of identifying the relevant comparison groups across waves.

In all specifications, all observations are weighted using DHS sampling weights, unless otherwise noted, and each specification includes controls for age, education, and district and division HIV rates.[20] Finally, each specification includes controls for urban-rural status, proximity to large and small health facilities, and each of these interacted with wave 2 to allow different trends.

The basic equation I estimate is:

$$Y_{ijt} = \beta_0 * Treat_j * Wave2_t + \beta_2 * Wave2_t + \gamma_j + \sum_{k=3}^{n} \beta_k * X_{kijt} + \epsilon_{ijt} \tag{1.12}$$

where Y_{ijt} is the outcome, $Treat_j$ is a binary variable that represents whether the respondent is located in an area in which ARVs were available before Wave 2, and γ_j is an area fixed effect. X_{ijt} is a vector of (n-3) individual-specific controls. Each wave surveys different villages, and therefore the definition of an area j cannot be a village. Each specification will define area differently.

In the preferred specifications, $Treat_j$ is defined as being within 8 kilometers of a facility with ARVs by 2008.[21] Because the same villages were not sampled across waves, the relevant

months pregnant at the time of the miscarriage) if not for the miscarriage are coded as pregnant. Results do not change if these are not coded as pregnant.

[19]Other individuals may learn about the presence of ARVs from those who have begun treatment either explicitly via word of mouth, or indirectly by observing health improvements of peers who are rumored to be HIV positive. These two channels of information could lead to the formation of different beliefs about HIV infections. In particular, indirect observation could erroneously signal that a cure is available. In the 2006 Uganda DHS, 34% of women who reported that they had heard of ARVs believed that they were a cure for HIV. As this belief is common, it is possible that behavioral responses to proximity to treatment could be driven by an over-estimate of the benefit of ARVs to those who are HIV positive. In this case, if individuals believe that ARVs are more effective than they are, they might respond more than they would have with accurate information.

[20]This is constructed using the DHS sample as this is the standard source of information about HIV rates. Each respondent is excluded from the estimate of the HIV prevalence in her area.

[21]Eight kilometers is chosen to maximize power as it is the closest distance to the median. This generates

comparison group across waves is not obvious. To address this, observations are matched across waves based on their locations using GPS locations to identify precise comparisons and construct a fixed effect analysis within pairs of neighboring survey clusters.

Each survey cluster in wave 2 is linked with the five closest survey clusters from wave 1.[22] For the analysis, each respondent from wave two is included five times and each observation from wave 1 is included as many times as it is matched. Any pair that is more than 100kms apart is dropped.

Using this expanded and matched sample, I estimate another difference in difference estimate with matched-pair fixed effects. Each specification includes a fixed effect for each matched-cluster pair. Each observation is additionally weighted by the minimum of the inverse of the distance and 1/8.[23] Pairs of cluster with different treatment status are dropped in the primary specification, but estimates including these as well are also presented, and do not generate noticeably different results. Dropping the unmatched pairs is comparable to excluding boundaries between areas in spatial analysis. The standard errors are clustered at the level of the survey cluster to correct for the duplication. I also report standard errors corrected for two-dimensional clustering following Cameron, Gelbach and Miller (2011). One dimension is a cluster from Wave 1 with all observations from Wave 2 with which it is matched, and the other dimension is the opposite. The standard errors are somewhat larger, but not substantially so. The coefficient of interest remains the interaction between $Wave2_t$ and $ARV\ Access_j$.

In a simpler specification, $Treat_j$ is defined as residing within a division in which at least one health facility provided ARVs by 2008. This specification includes division fixed effects and standard errors clustered at the level of the division.[24,25] While divisions can be large, this measure of proximity may reflect reality in that individuals are likely to visit the center of their division for other business, even if they do not live as close. Therefore it is logical that the relevant proximity that would determine the spread of information about a new HIV treatment could be within the same district. The geographic distribution of treated divisions is shown in Figure 1.2. One weakness of this specification is that observations from divisions with clusters in only one round do not contribute the estimates, so information is lost, which is why the matched specification is preferred.

Robustness is verified using multiple age cut-offs, and results are also reported separately for those married and unmarried. The theoretical framework suggests a change in behavior among those who are HIV negative. The analysis that follows includes a very small fraction of respondents who tested positive for HIV. The results are robust to excluding this group.

balance between the treatment and control groups that maximizes the precision of the estimates. This distance (approximately 5 miles) is also a reasonable distance to walk for frequent medical care. For robustness, the analysis is repeated using different distance cut-offs with nearly identical results.

[22]Because the locations of villages is jittered and some villages may be sampled twice, it is possible that some of these matched pairs are truly taken from the same villages at two points in time.

[23]This weighting scheme is used in place of the inverse distance so as not to overweight extremely small distances. Because of the jittered data, these distances are not likely to be precise at this level.

[24]During the time between waves, administrative boundaries have shifted. For consistency, I use current borders and place observations within them using their GPS locations.

[25]Due to jittering, 11 clusters were placed outside of the borders of Kenya. These observations were manually linked with the closest administrative division within the country so that they could be included in this analysis.

All estimates include controls for age, education, district and division HIV prevalence, urban-rural status, and proximity to other health facilities, along with survey wave and location fixed effects as described.

The primary assumption to justify the difference in difference specification is that the trends in the treatment and control areas would have been the same in the absence of treatment. Figure 1.4 plots pregnancy rates in treated and control areas (defined using the 8km distance threshold), before 2003, based on the birth registry in the DHS data. While the levels are not the same, the trends are clearly similar, and we cannot reject that the two curves are parallel.

1.5 Results

The main results are reported in Panel A of Table 1.3. Columns 1 and 2 present the results using the specifications with matched clusters of observations. In Column 1, this estimation includes all matches, and Column 2 excludes the pairs with different treatment status from the analysis. The treatment effect is the coefficient on the interaction term, reported in the first row. This shows a treatment effect of 6.7 percentage points. Column 3 presents the specification in which treatment is defined as having a facility with ARVs in the same division, showing a treatment effect of 9.5 percentage points. In all three specifications, the coefficient of interest is positive and statistically significantly different from zero.

Panels B and C of Table 1.3 repeat the same estimation, using whether the respondent reports that she has had sex in the last 4 weeks as the outcome. In Panel C, the outcome is reporting having had sex in the last 4 weeks and reporting having not used a condom with the most recent sexual partner. In the first and second columns, the treatment effect is measured to be approximately 5 percentage points. While sexual activity would need to change by a larger magnitude to generate the observed changes in pregnancy rates, the lower estimated treatment effects could reflect attenuation from noise resulting from misreporting. In the third column, the coefficient of interest is insignificant, but the point estimate and standard error are both large, so a substantial increase cannot be rejected.

Fertility preferences

Changes in pregnancy rates could also reflect differences in fertility preferences, questioning the applicability of the proposed theory of risk-taking to explain the observed results. Panel A of Table 1.4 estimates the impact on other measures of fertility preferences or access to family planning, using the matched specification with only pairs with similar treatment status (Estimation Strategy 3). Column 1 shows the impact on having been visited by a family planning worker as a test of whether the introduction of ARVs also brought changes in the provision of broader reproductive health services. This coefficient is negative, small, and insignificant. The second column estimates the impact on the stated ideal number of children. Treatment areas - which were observed to have had relatively higher conditional fertility rates - reported lower numbers of ideal children. Columns 3 and 4 find no impacts on the use of birth control, conditional and not conditional on having had sex respectively.

Another way to test whether the impacts on pregnancy reflect changes in general fertility preferences is to look at which segment of the population changes their behavior. If ARVs changed fertility preferences, then we would expect to see a change in fertility among

those who are married at least as strongly as among those who are not married. Panel B of Table 1.4 repeats the main analysis using different subgroups. Column 1 includes married women, and column 2 includes women who have been married for at least 1 year. Column 3 includes women who report that they are cohabiting. In each of these three specifications, the estimated treatment effect is either negative or extremely small and insignificant. However, Column 4 includes those who are unmarried and over 25 (in order to have a completely distinct population from those in the previous estimates), including those who never married or are divorced or widowed. In this specification, the treatment effect estimated is 4.6 percentage points, similar to that estimated for young women. Selecting on these subgroups is problematic because the criteria for selection are potentially endogenous and could themselves be responses to ARV access. Still, these estimates suggest that the measured differences are less likely to be a reflection of changes in fertility preferences, and they may reflect differences in risk preferences with regards to unprotected sex among populations with the possibility for marginal behavior changes.

Columns 1 and 2 of Table 1.6 present estimates of the impact on unwanted pregnancies, as these are more likely to reflect changes in risk-taking rather than fertility preferences. I code pregnancies as unwanted if the respondent reports that she did not want to become pregnant or did not want to become pregnant at that time. For those who have recently miscarried but would have been pregnant otherwise, respondents are not asked whether they wanted the pregnancy. I code all pregnancies resulting in miscarriage as unwanted, and the results are nearly identical if these are all coded as wanted. Rates of reported unwanted pregnancies are substantially lower than for all pregnancies, and so the estimated impacts are correspondingly smaller, but still positive and substantial.

HIV testing

As discussed earlier, Wilson (2011) demonstrates that demand for HIV testing is likely to increase with ARV access. This increase in testing could facilitate partner sorting based on HIV status or *sero-sorting*. This presents an alternative channel by which ARV access increases testing which facilitates sero-sorting, which increases pregnancies among those who know their partners status and thus are not putting themselves at risk of HIV infection. While this could be part of the story, there is evidence that it is not the entire story. First, in this sample, even in the second wave, only 27 percent of those in areas with ARVs had been tested, while 21 percent of those in control areas had been tested. Of those who were tested in treatment areas in wave 2, only one third (or 9 percent of the entire group) had been tested more than one year before the survey. Columns 3 and 4 of Table 1.6 repeat the main analysis excluding those who had been tested at least one year before the survey, and the results remain the same. While sero-sorting may marginally contribute to the increase in pregnancy among young women, it cannot explain the observed relative increase in risky behavior in areas that received ARVs.

HIV testing could also change beliefs about own HIV status, as many individuals overestimate the probability that they are infected. If this is the case, as ARV access encourages testing more people will believe that they are HIV negative. Following Gong (2011), this would predict a reduction in risk-taking as those who believe they are likely to be positive and find they are negative were demonstrated to respond by taking fewer risks, and so this cannot be driving the results. On the other hand, if this also changes beliefs about the prevalence of HIV in the general population, this becomes more complicated. Without detailed

information about beliefs, this is beyond the scope of this paper.[26]

The threshold of 8 kilometers was chosen because it is near the median in order to maximize power, but - like any other distance cutoff - it is somewhat arbitrary. Panel A of Table 1.7 allows the distance threshold to vary from 8-12 kilometers. Each column repeats the analysis of the first column of Panel A of Table 1.3 with pregnancy as the outcome using a different distance cut-off. The results are remarkably consistent across these specifications. The sample sizes varies somewhat because of the restriction that matched pairs have the same treatment status. Some misclassification is inevitable as any distance cut-off will necessarily put some individuals who know about treatment outside of the circle while including others who do not know about it within it. However, this demonstrates that the particular choice of the threshold does not determine the estimated results.

The age cut-off can also be varied to show that there are consistent results using alternative age thresholds. While the main cut-off restricts the analysis to teenagers, a demographic that is of particular interest in research on changes in fertility behavior, others are possible. For example, the majority of those aged 21 and under do not have children, while those above are more likely than not to have had a child. The majority of those 22 and under do not report that they are cohabiting and the majority of those 23 and under do not report that they are married. Panel B of Table 1.7 repeats the analysis from Column 3 of Table 1.3 varying the age cutoff from 19 to 24, and Panel C of Table 1.7 repeats the analysis from Column 3 of Panel A of Table 1.3 using the administrative area to determine treatment status. In both tables, the results are reasonably consistent, although the estimated treatment effect declines as the threshold increases. The increase in age increases the proportion of the sample that is already married, cohabiting, or otherwise in a stable partnership, and thus unlikely to respond to changes in risk of unprotected sex, and this is likely to generate the decline in the estimated effect.

1.6 Simulation

The introduction of antiretroviral drugs could influence the spread of HIV both through changing behavior and through biological channels - reducing infectiousness of those on treatment and keeping more people who are HIV positive alive. This is formalized in Section 1.2, demonstrating how the sign of the impact of ARVs on new infections is ambiguous and depends on behavior.

The empirical analysis above showed a relative increase in risk-taking among those with access to antiretroviral treatment. This can directly increase the rate of new infections by increasing those who put themselves at risk. However, it also can indirectly decrease the rate of new infections as the increase in A_1 means that a larger fraction of the pool of potential sexual partners is HIV negative, decreasing the risk of infection for those who engage, p. This is formally demonstrated by Kremer (1996).

In addition, the reduction in transmission risk from treatment, q, can outweigh a substantial change in behavior among those who are negative so that the rate of new infections will decline with treatment. It bears mentioning that beyond the impact on new infections,

[26]For more information about changes in beliefs as a result of ARV access, see Baranov, Bennett and Kohler (2012).

ARV access has large and important welfare impacts for those who are infected and receive treatment.

In practice, the effect of D (the level of ARV provision) on behavior is likely to be non-linear with substantially larger effects on behavior when the marginal person put on treatment is sicker. The benefit to an individual who is HIV positive of being on treatment is high when he or she has a low CD4 count, which means being close to AIDS onset and opportunistic infections. However, especially given the toxicity and unpleasant side-effects, earlier treatment is not likely to provide a significant additional benefit to the individual. Thus while access to treatment provided to individuals with a CD4 count below 200 (which was previously the WHO recommended threshold) can generate the observed difference in behavior, the behavioral response is not likely to grow as the CD4 count threshold increases. However, the change in this threshold will change the probability of infection as more infected individuals are put on treatment and present a lower transmission probability.[27]

Based on the reasoning above, a low level of ARV access could change behavior but not lead to a significant reduction in infectiousness, while a very high level in which treatment is available upon diagnosis of HIV infection would reduce incidence of HIV. This is outlined in Over et al. (2006) and Granich et al. (2009) who propose beginning treatment immediately after a positive HIV test.

This will be demonstrated via simulation. Recall

$$I = A_1 * p$$

where I is the rate of new infections, A_1 is the fraction of the negative population that has unprotected sex, and p is the likelihood of transmission conditional on unprotected sex.

This probability can be written as:

$$p = r * \frac{A_2 N_2 + A_3 N_3 q}{A_1 N_1 + A_2 N_2 + A_3 N_3}$$

where N_j is the size of group j, A_j is the proportion of each group that has unprotected sex, r is the transmission risk from sex with a Type 2 individual, and $r * q$ is the transmission risk from sex with a Type 3 individual. The simulation will use available estimates of each of these parameters to estimate the impact of drugs on new infections. For clarification, treatment changes A_1, N_2, and N_3. The assumptions used in the simulation are summarized in Table 1.5.

As described above, treatment changes behavior most at the low end, but would not be expected to change dramatically as access is available to anyone with a sufficiently low CD4-count, while the impact on transmission rates continues as treatment is provided to those based on higher CD4 thresholds. Based on Williams et al (2006), if the CD4 count threshold is set at 200, then 17% of those who are HIV positive will receive treatment. This number climbs to 44% if the threshold is 350 and 67% if the threshold is 500. For simplicity,

[27]WHO changed the recommended CD4 count threshold to determine ARV eligibility from 200 to 350, however most countries in Sub-Saharan Africa have not reached full coverage even with the lower threshold due to a lack of supplies. Rwanda is one exception, reporting nearly 100 percent coverage of those eligible, and experimenting with using 500 as a threshold for those in sero-discordant couples to reduce the likelihood of transmission to the uninfected partner.

I assume that below 17%, treatment is given to a fraction of those who need it and behavior changes for this fraction of the negative population. Above this threshold, behavior change is constant, at the level estimated in the empirical analysis. This assumed relationship between the fraction positive on treatment and the fraction negative who have sex is demonstrated in Figure 1.5.

I simulate new infection rates at all levels of drug provision up to 67%. This is done using 10,000 individuals. First, HIV status is assigned, then some are assigned to treatment based on the level of distribution. Death rates determine survival, and some choose to have unprotected sex. Of those who choose to, they are matched randomly. Some become infected. This is repeated for each percentage on treatment from 0-67% 500 times with and without behavior change, and with q equal to 1, 0.5 and 0.04.

Figures 1.6 and 1.7 present the estimated infection rates. Figure 1.6a assumes that there is no behavioral response and no reduction in transmission, and clearly, there is nearly no difference in new infection rates, except for a moderate increase explained by keeping more people who are infected alive. Figure 1.6b also presents estimates with no reduction in transmission, but with a change in behavior. This presents a much larger increase in infection rates. Figure 1.7a presents infection rates for different levels of treatment distribution if the reduction in transmission probability from ARVs is substantial ($q = 0.04$). Here, there is a slight jump in infection rates when behavior changes (at the CD4 count threshold of 200), but there is a substantial decline in infection rates that outweighs this. Figure 1.7b uses $q = 0.5$ to show the impact of ARV provision if the reduction in transmission is more modest. In this case, the increase in infection due to behavior change is outweighed only if a sufficient fraction of the population is put on treatment. This suggests that provision of treatment can decrease infection rates, but that overcoming a behavioral response depends on reaching a sufficient threshold.

1.7 Conclusion

Previous models of the impacts of ARVs insufficiently acknowledged the importance of behavior change in shaping HIV incidence. With the absence of evidence about the magnitude or sign of this behavioral response, even the direction of the response could only be guessed. However, taking this response seriously is necessary for credibly evaluating drug provision to inform developing country governments and international donors as they weigh competing demands on tight budgets. This paper fills two prominent holes in the existing literature on HIV treatment provision in Sub-Saharan Africa: First, it provides the first causally identified estimates of the change in risky behavior due to treatment access in the context of a generalized epidemic. Second, it shows how these estimates work with existing medical evidence about the mechanical effects of ARVs to determine the predicted impacts of treatment provision on new HIV infections.

Using an original dataset that combines administrative records of the roll-out of treatment facilities in Kenya with two national population surveys, I estimate a substantial increase in risk-taking in response to treatment access. Among young women, this demonstrates an increase in pregnancies of 82% and an increase in self-reported sexual behavior of 40%. Identifying this response is crucial to estimating the impact of ARVs on the course of the

HIV epidemic. Incorporating the behavioral response into a simulated model of the impact of different levels of ARV provision demonstrates that treatment provision can reduce new infection rates, even with the substantial increase in risk-taking estimated in the empirical section of the paper.

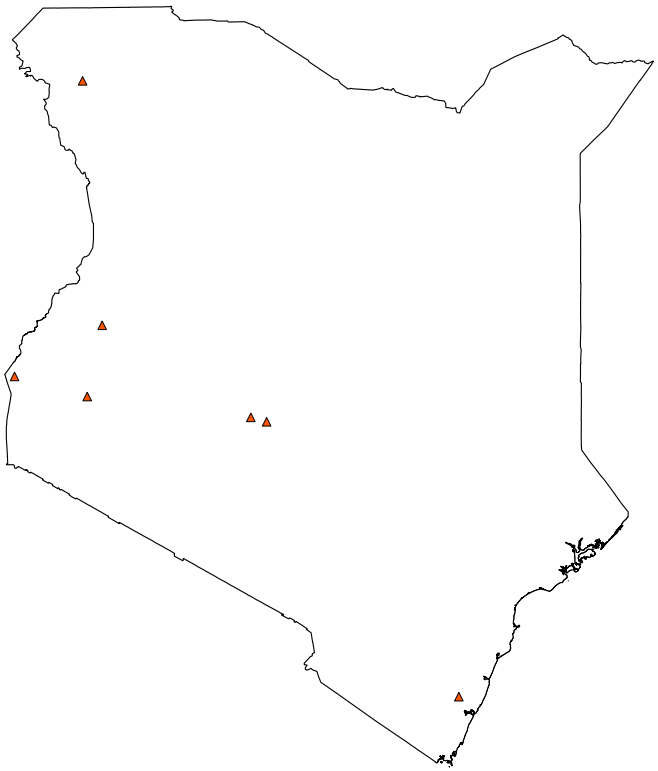
Like any study with data from a single country, the question of generalizability remains. Future work will apply the same method of analysis in Uganda and Rwanda, combining administrative records of ARV distribution, which I have already collected, with recently released DHS data from these countries. While these results diverge from previous studies in Sub-Saharan Africa that do not find significant changes in risk-taking in response to information about HIV risk (e.g.: Godlonton, Munthali and Thornton (2011), Oster (2012), Wilson, Xiong and Mattson (2012)), they match evidence of behavioral responses to ARV provision among gay men in the US (Mechoulan (2007), Papageorge (2012)). While previous changes in the risk environment were generated by variation in the likelihood of infection, ARVs change the costs of infection. As the likelihood of infection from a single encounter is low, perhaps the changes in probabilities are not easily understood or perceived, whereas a change in life expectancy and the cost of infection is more salient.

While this paper provides some evidence of the extent to which risky sexual behavior responds to changes in the cost of HIV infection, more work remains to be done to assess the generalizability of these results and variation in responses among different populations. Hopefully future assessments of proposed policy changes regarding HIV treatment provision will acknowledge the potential strength and importance of behavioral responses.

1.8 Figures

Figure 1.1: ARV distribution sites in Kenya

(a) 2004



(b) 2009

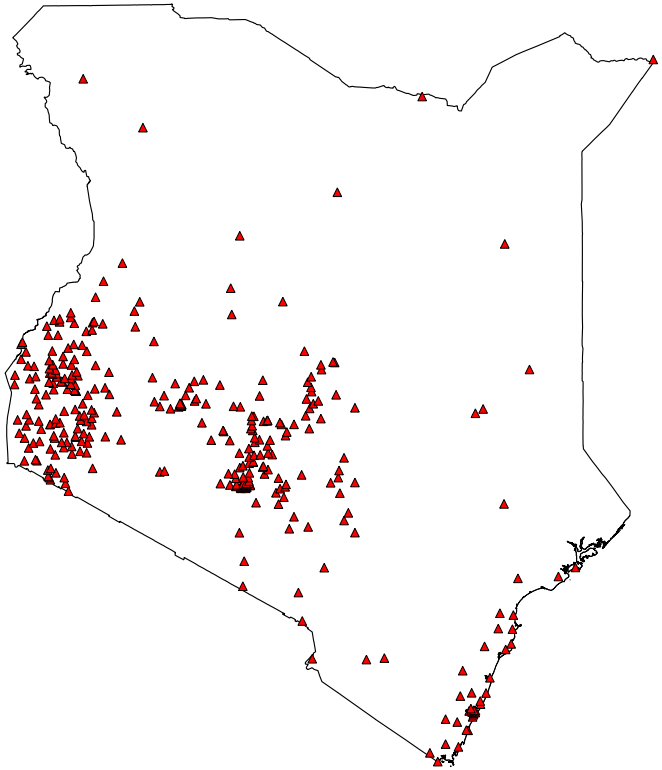


Figure 1.2: Treated divisions

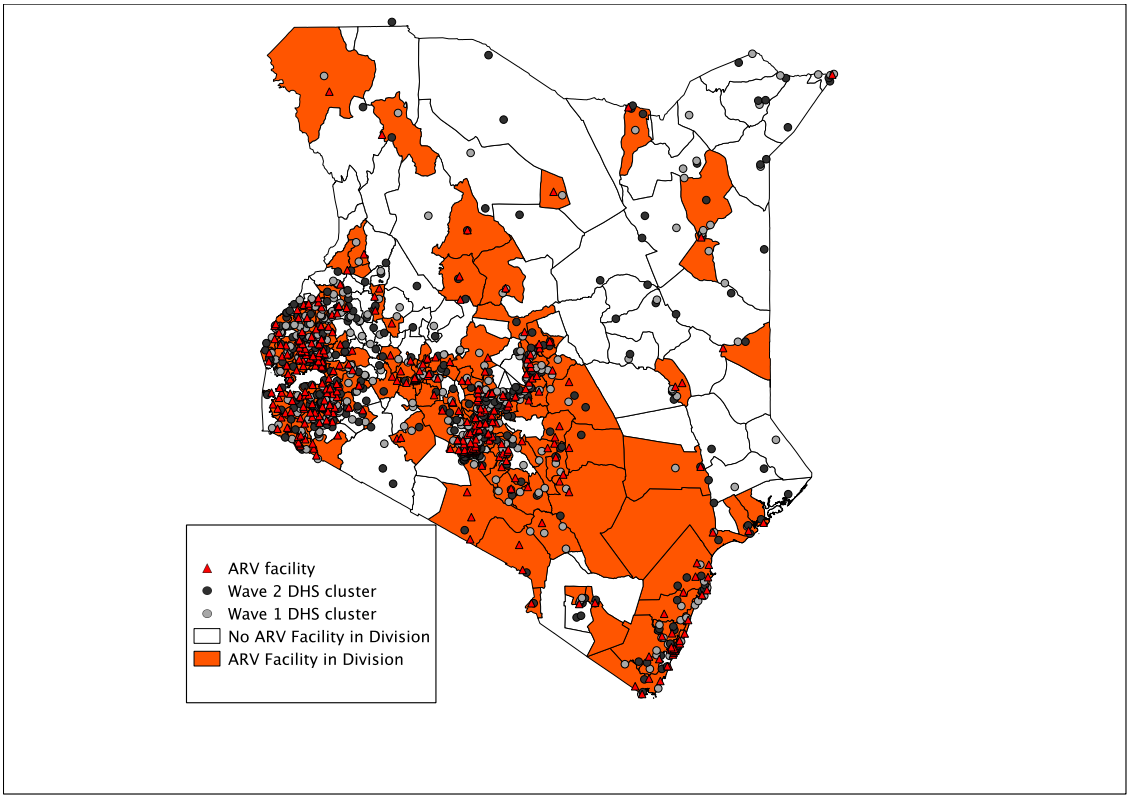


Figure 1.3: DHS clusters in Kenya, 2003, 2008/2009

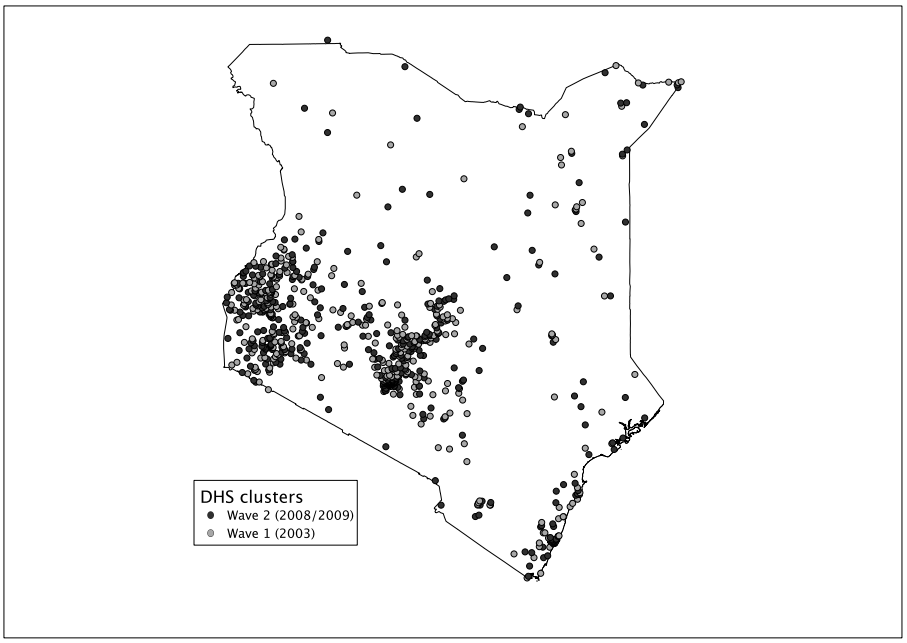


Figure 1.4: Parallel trends: Pregnancy rates before 2003

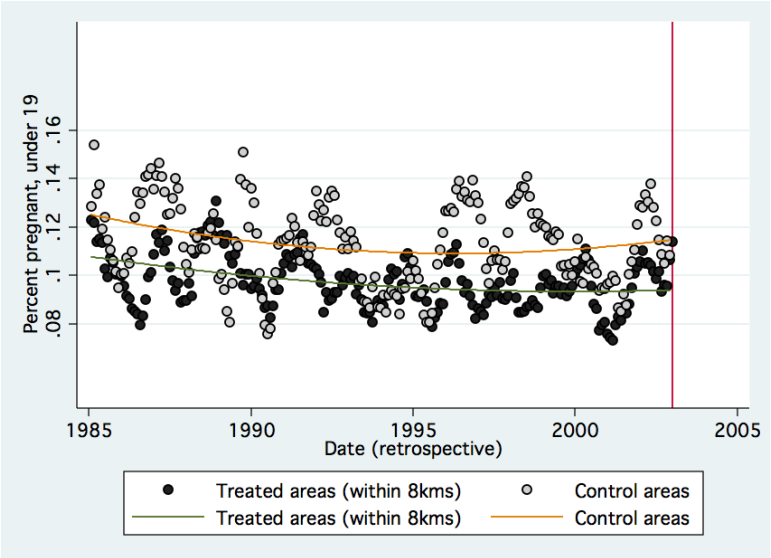


Figure 1.5: Simulation assumption of behavior change

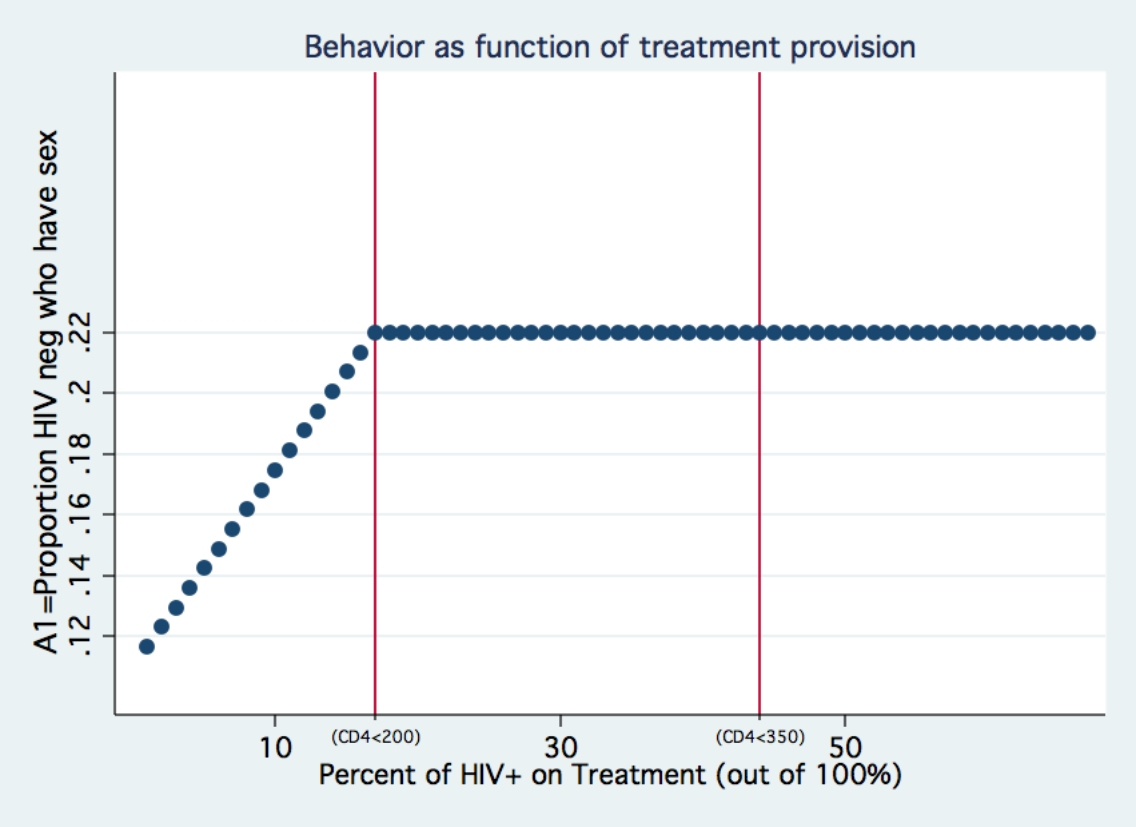
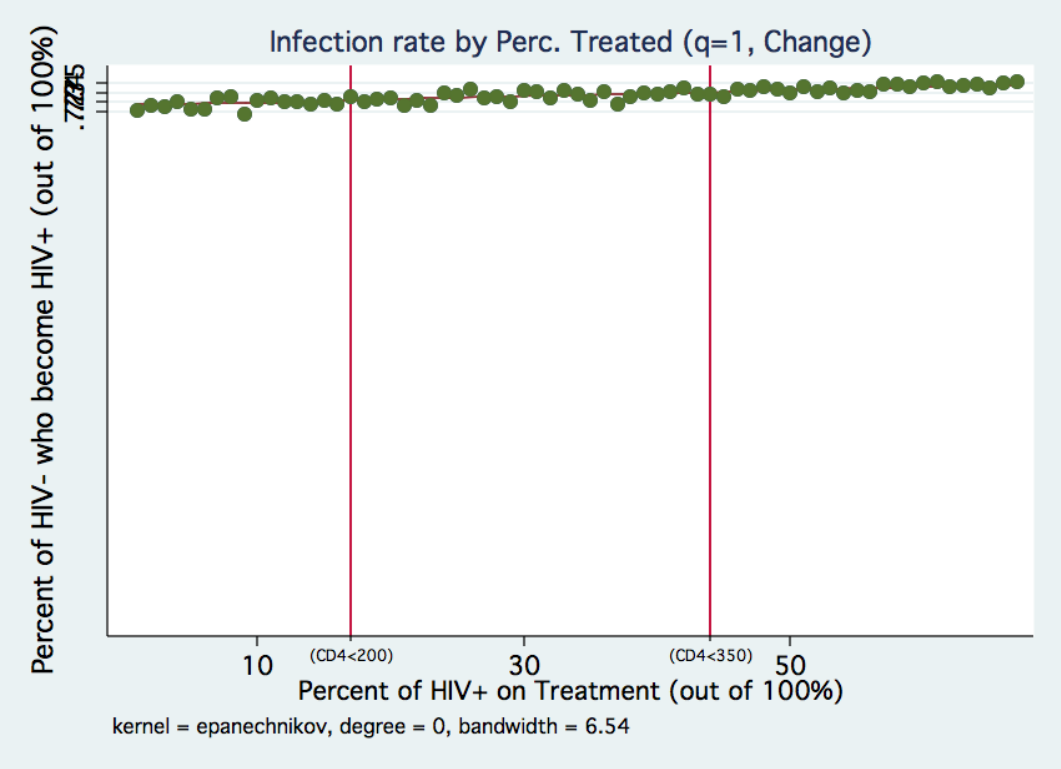


Figure 1.6: No reduction in transmission probability (q=1)

(a) No behavior change



(b) Behavior change

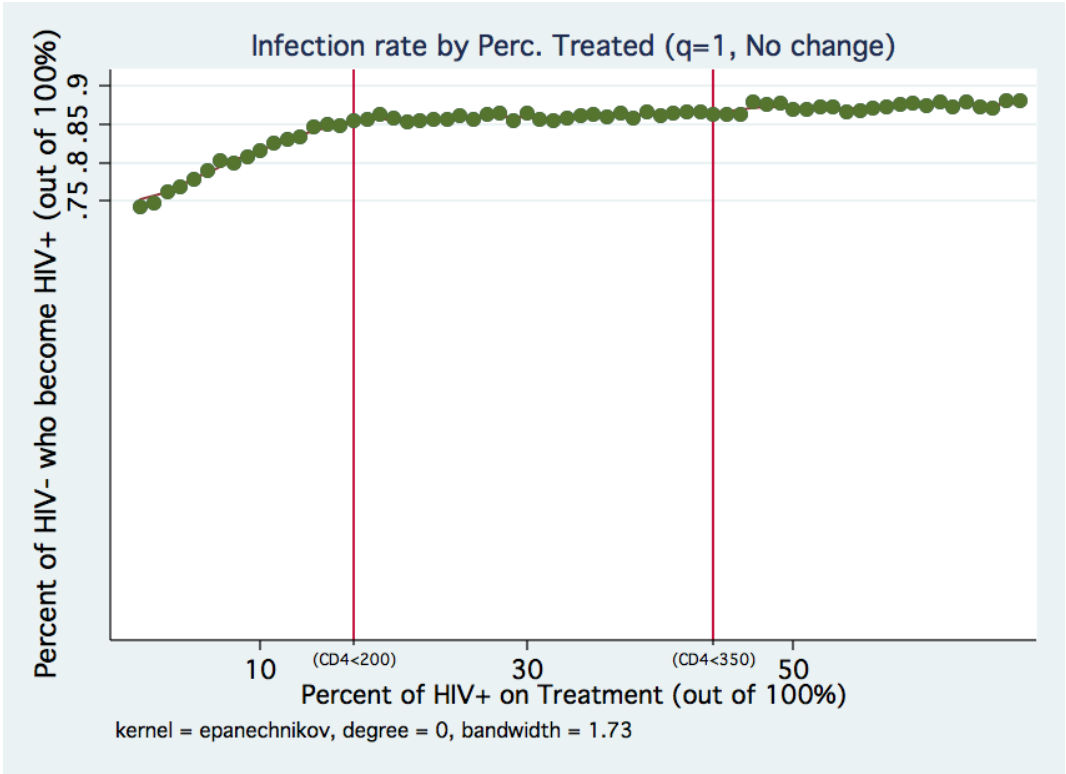
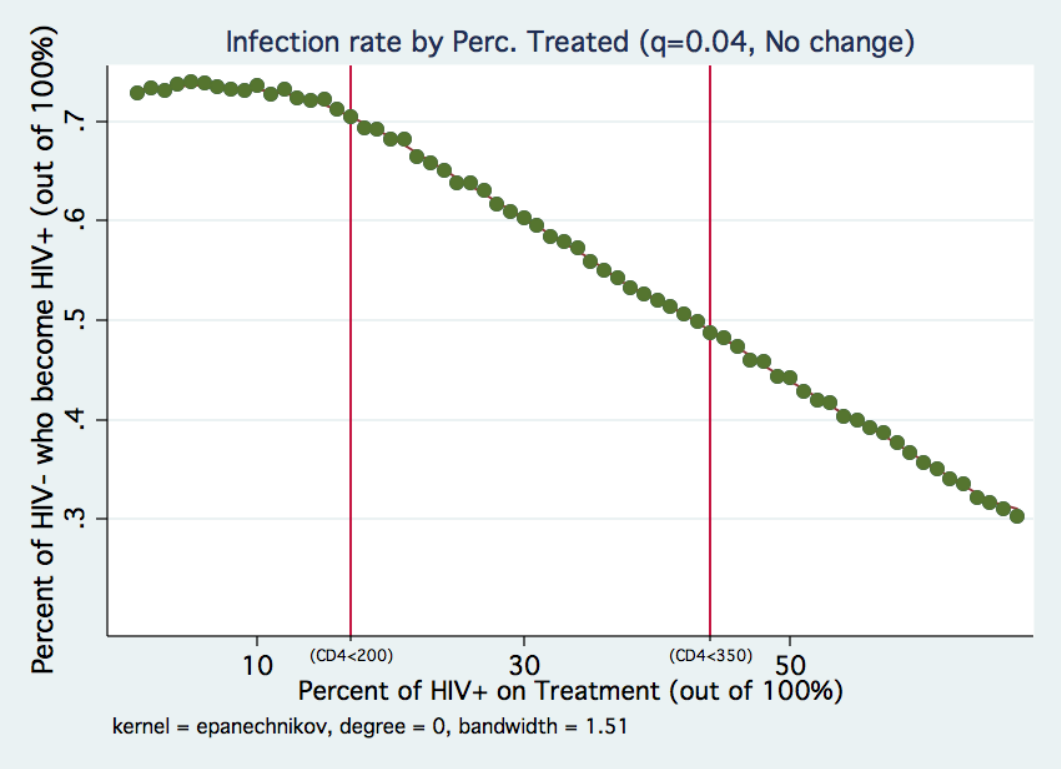
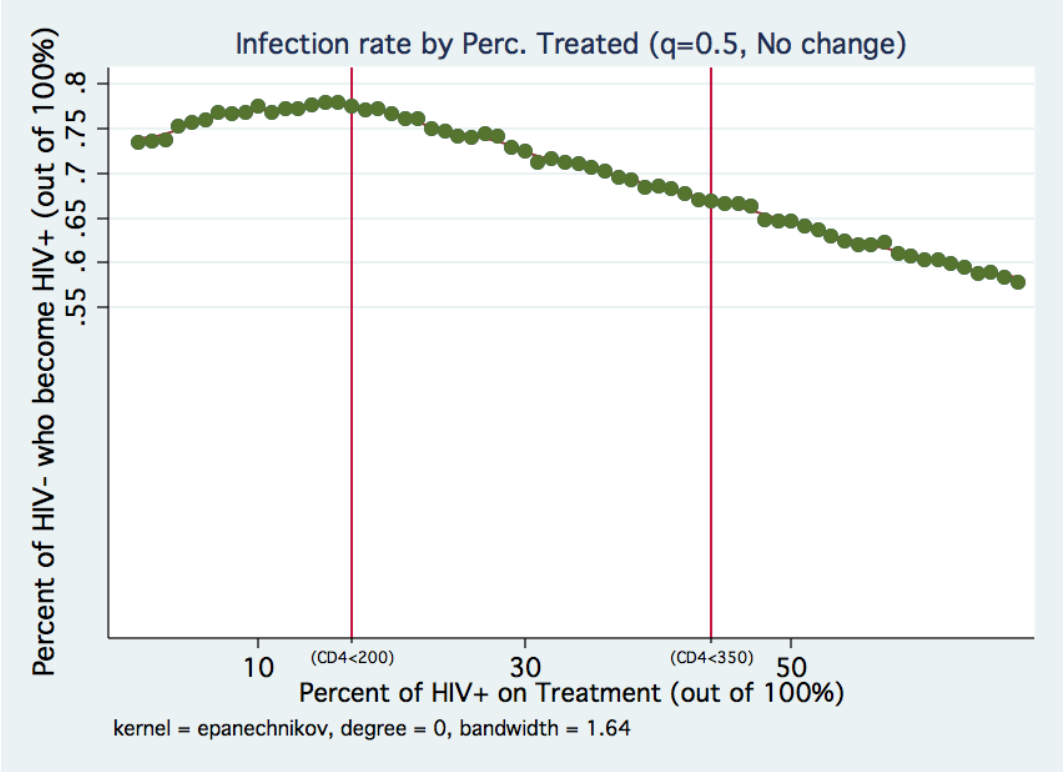


Figure 1.7: Behavior change and reductions in transmission probability

(a) Reduction in transmission probability of 96% (q=0.04)



(b) Reduction in transmission probability of 50% (q=0.5)



1.9 Tables

Table 1.1: Summary of ARV roll-out, HIV prevalence, and survey timing

	Number Facilities	HIV Prevalence	DHS survey	
Year	with ARVs	(WHO)	Female Respondents	Clusters
2003	1	7.5	8,195	400
2004	7	7.1		
2005	153	6.8		
2006	188	6.6		
2007	263	6.4		
2008	336	6.3	8,444	398
2009	392	6.3		
2010	610			

Note: Facilities counted as distinct only if in different locations.

Table 1.2: Summary Statistics 1

	2003		2008/2009	
	No ARVs in 8kms by 2008	ARVs in 8kms by 2008	No ARVs in 8kms by 2008	ARVs in 8kms by 2008
HIV positive	.014	.037	.01	.059
	(.119)	(.189)	(.099)	(.236)
Years of education	5.978	6.867	6.979	7.764
	(2.853)	(2.515)	(2.751)	(2.285)
Married	.151	.062	.062	.082
	(.358)	(.241)	(.241)	(.275)
Heard of AIDS	.966	.995	.979	.994
	(.183)	(.071)	(.145)	(.079)
Knows someone who has or died of AIDS	.656	.669	.614	.72
	(.476)	(.471)	(.487)	(.449)
Ever been tested for AIDS	.033	.054	.21	.282
	(.179)	(.225)	(.408)	(.451)
Ever had sex	.358	.351	.296	.331
	(.48)	(.477)	(.457)	(.471)
Had sex in the last 4 weeks	.144	.115	.066	.118
	(.351)	(.319)	(.249)	(.323)
Currently Pregnant/Miscarried	.083	.04	.035	.047
	(.276)	(.195)	(.185)	(.212)
Current unwanted pregnancy/miscarriage	.038	.024	.012	.035
	(.19)	(.153)	(.108)	(.184)
Ideal number of children	3.808	3.318	3.636	3.174
	(2.449)	(1.741)	(1.927)	(1.585)
Used any birth control method	.034	.055	.032	.06
	(.181)	(.228)	(.177)	(.238)
Used any birth control if had sex	.082	.141	.11	.18
	(.275)	(.349)	(.314)	(.385)
Has at least two sexual partners	.009	.021	.008	.011
	(.095)	(.144)	(.087)	(.106)
Had any STD in last 12 months	.002	.006	.003	.003
	(.047)	(.079)	(.055)	(.056)
Had STD symptoms in last 12 mos.	.009	.014	.005	.017
	(.096)	(.118)	(.074)	(.129)

Note: Standard deviations in parentheses. Includes women ages 15-18. Excludes areas with ARVs before 2004.

Table 1.3: Impacts of ARV access on pregnancy, self-reported sexual activity

	VARIABLES	(1) Matched Same Treatment Status	(2) Matched	(3) ART in Division
Panel A: Currently Pregnant	ARV Access*Wave2	.064 ***	.067 ***	.095 ***
		(.02)	(.019)	(.033)
		[.033]	[.033]	
	Observations	11,391	7,538	2,494
	Clusters	621	583	207
Panel B: Sex in the last 4 weeks	ARV Access*Wave2	.057 ***	.057 ***	.027
		(.024)	(.024)	(.041)
		[.042]	[.042]	
	Observations	11,391	7,538	2,494
	Clusters	621	583	207
Panel C: Unprotected sex in the last 4 weeks	ARV Access*Wave2	.048 **	.048 **	.03
		(.022)	(.022)	(.041)
		[.037]	[.037]	
	Observations	11,391	7,538	2,494
	Clusters	621	583	207

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large and small health facilities within 10kms, and each of these location characteristics interacted with Wave 2. Columns 1 and 2 include pair fixed effects with standard errors clustered at the level of the survey cluster. Two-way clustering adjusted standard errors, following Cameron et al (2006) are reported in square brackets. All estimates are weighted using DHS sampling weights. Estimates in columns 1 and 2 are additionally weighted by the DHS sampling weights multiplied by the minimum of 1/8 and the inverse of the distance between the pair. Column 3 includes division fixed effects and standard errors, clustered at the division level.

Table 1.4: Impacts of ARV access on fertility preferences, pregnancy in alternative subsets

		(1)	(2)	(3)	(4)
		Visited by FP worker	Ideal number of children	Used any birth control, if had sex	Used any birth control
Panel A: Fertility Preferences	ARV Access*Wave2	.007	-.207*	.017	.009
		(.011)	(.123)	(.057)	(.019)
		[.018]	[.219]	[.161]	[.033]
	Observations	7521	7538	2534	7538
	Clusters	583	583	402	583
		(1)	(2)	(3)	(4)
		Pregnant	Pregnant	Pregnant	Pregnant
Panel B: Alternate Subsets	ARV Access*Wave2	-.014	-.016	-.003	.059***
		(.013)	(.013)	(.012)	(.022)
		[.021]	[.021]	[.021]	[.04]
	Observations	23343	22376	25523	6224
	Clusters	620	620	620	561
	Subset:	Married	Married one year or more	Cohabiting	Unmarried Over25

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large and small health facilities within 10kms, each of these location characteristics interacted with Wave 2, and pair fixed effects. Standard errors in parentheses are clustered at the level of the survey cluster. Two-way clustering adjusted standard errors, following Cameron et al (2006) are reported in square brackets. All estimates are weighted using DHS sampling weights multiplied by the minimum of 1/8 and the inverse of the distance between the pair.

Table 1.5: Simulation Assumptions

Parameter	*Value*	*Notes*
r (transmission probability)	0.23	(representing one year)
q (reduction in transmission probability with ARVs)	0.04, 0.5, 1	0.04 represents estimates from Cohen et al (2011) 0.5 represents the lowest end of medical estimates 1 represents no reduction
d_1 (death rate among HIV negative)	0.027	Average mortality for 15-19 year-olds in Kenya between 2000 and 2005: World *Population Prospects: The 2010 Revision* UN Department of Economic and Social Affairs, Population Division (2011)
d_2 (death rate among HIV positive, untreated)	0.12	
d_3(death rate among HIV positive, treated)	0.06	
A_2 (proportion of positive untreated who have unprotected sex)	0.37	Fraction of HIV positive DHS respondents who reported having had sex in previous four weeks in untreated areas
A_3 (proportion of positive and treated who have unprotected sex)	0.33	Fraction of HIV positive DHS respondents who reported having had sex in previous four weeks in treated areas
A_1 without ARVs	0.11	Assuming: pregnancy lasts 9 months, individuals have sex twice per week, the pregnancy rate when drugs are not available is 0.6, the likelihood of becoming pregnant from unprotected sex once is 0.01: A_1 (without ARVs)= $\frac{0.07}{1-(1-0.01)^7 8} = 0.12$
A_1 with ARVs	0.11	With a pregnancy rate when drugs are available of 0.12: A_1 (with ARVs)= $\frac{0.12}{1-(1-0.01)^7 8} = 0.22$

Table 1.6: Robustness Checks

VARIABLES	(1) ART in Division Unwanted Preg	(2) Matched Same Status Unwanted Preg	(3) ART in Division Pregnant (Untested)	(4) Matched Same Status Pregnant (Untested)
ARV Access*Wave2	0.0485**	0.0134	0.0785**	0.0589***
	(0.0245)	(0.0114)	(0.0338)	(0.0184)
Observations	2,494	7,538	2,367	7,100
R-squared	0.105	0.188	0.141	0.225
Clusters	207	583	207	579

Note: All estimates include controls for age and education, district and division HIV prevalences, urban-rural status, the presence of large and small health facilities within 10kms, and each of these location characteristics interacted with wave 2. Columns 1 and 3 define treatment as an ARV provision facility in the same division, and they include division fixed effects and standard errors, clustered at the division level. Columns 2 and 4 define treatment by distance and include include pair fixed effects with standard errors clustered at the level of the survey cluster. The dependent variable in columns 1 and 2 is current unwanted pregnancy. Columns 3 and 4 restrict the sample to those who have not been tested for HIV in the previous 12 months. All estimates are weighted using DHS sampling weights. Estimates in columns 2 and 4 are additionally weighted by the DHS sampling weights multiplied by the minimum of 1/8 and the inverse of the distance between the pair.

Table 1.7: Robustness: Impacts of ARV access on pregnancy, Alternative specifications

Panel A: Treatment defined as within fixed distance, Varying cutoff distance

	(1) Cutoff: 6km	(2) Cutoff: 7km	(3) Cutoff: 8km	(4) Cutoff: 9km	(5) Cutoff: 10km	(6) Cutoff: 11km
ARV Access*Wave2	.06***	.054***	.064***	.055***	.061***	.066***
	(.019)	(.02)	(.02)	(.021)	(.022)	(.025)
	[.032]	[.033]	[.033]	[.034]	[.036]	[.041]
Observations	7,422	7,442	7,538	7,947	8,050	8,242
Clusters	588	585	583	589	585	576

Panel B: Treatment defined as within 8kms, Matched specification, Varying ages

	(1) Under 18	(2) Under 19	(3) Under 20	(4) Under 21	(5) Under 22	(6) Under 23
ARV Access*Wave2	.035**	.064***	.055***	.049***	.036**	.021
	(.016)	(.02)	(.018)	(.018)	(.016)	(.015)
	[.027]	[.033]	[.029]	[.03]	[.026]	[.026]
Observations	5,653	7,538	9,313	11,266	12,744	14,464
Clusters	561	583	602	612	615	617

Panel C: Treatment defined as within same division, Varying ages

	(1) Under 18	(2) Under 19	(3) Under 20	(4) Under 21	(5) Under 22	(6) Under 23
ARV Access*Wave2	.061**	.095***	.105***	.083***	.073***	.064**
	(.03)	(.033)	(.03)	(.029)	(.028)	(.031)
Observations	1,867	2,494	3,077	3,728	4,205	4,766
Clusters	207	207	207	208	208	209

Note: In all specifications, the outcome is an indicator for whether the respondent is currently pregnant All estimates include controls for age and education, district and division HIV prevalences urban-rural status, the presence of large and small health facilities within 10kms, and each of these location characteristics interacted with Wave 2. Panels A and B include pair fixed effects with standard errors clustered at the level of the survey cluster. Two-way clustering adjusted standard errors, following Cameron et al (2006) are reported in square brackets. Estimates in panels A and B are weighted by DHS sampling weights multiplied by the minimum of 1/8 and the inverse of the distance between the pair. Estimates in panel C are weighted by DHS sampling weights and include division fixed effects and standard errors, clustered at the division level.

Chapter 2

Corruption and the Effectiveness of Imported Antiretroviral Drugs in Averting HIV Deaths

2.1 Introduction

Today antiretroviral drugs are widely available in sub-Saharan Africa, with 8 million people receiving treatment in 2011 according to the World Health Organization. Until the last decade, this level of provision was considered inconceivable as the drugs were prohibitively expensive, and this enormous expansion in access has been credited with extending the lives of millions of people across the continent. At the same time, corruption in governments is associated with inefficient distribution of public goods, and this could limit the effectiveness of imported drugs in saving lives if the drugs do not reach the intended clinics or individual recipients, or if they are distributed with insufficient guidance.

This paper addresses the role of corruption in determining the effectiveness of antiretroviral drugs in reducing HIV mortality in sub-Saharan Africa. This is first done using a cross-country analysis comparing the impact of imported drugs on HIV deaths across countries with different levels of corruption. This is done using an original panel dataset of countries in sub-Saharan Africa from 2000-2007. This dataset combines standard measures of corruption used in economics and political science, information about HIV prevalence and deaths, and records of the quantities of antiretroviral drugs imported into each country. Using year and country fixed effects, this data provides evidence that HIV deaths are reduced less in corrupt countries given the same quantity of medicine, and the effect is even larger if the relevant quantity of drugs is measured in dollars spent.

There are many channels through which corruption could influence the effectiveness of health investments. For example, drugs can be purchased and then diverted either outside the country or within the country. The supply chain can fail if governments with higher levels of corruption are generally less capable of delivering public goods. Additionally, corruption within a government can facilitate targeting of public goods, not based on need, but based on political or other motivations.

Diversion of drugs could happen if drugs purchased by governments are resold. This could

be particularly lucrative in sub-Saharan Africa for two reasons. First, in nearly all countries of sub-Saharan Africa, supply is not nearly sufficient to meet demand and so treatment is rationed. This makes resale valuable because some of those excluded are likely to be willing to pay for the treatment. Second, because of international agreements with pharmaceutical companies, ARVs are sold at an enormous discount to governments and NGOs working in many countries in sub-Saharan Africa. This variation in price between different countries creates a substantial opportunity for arbitrage.

Such diversions prevent the drugs from reaching those who need them most, and they may take them out of the country entirely. It should be noted that if these drugs are sold to others within the same country, then a change in allocation may not reduce the overall reduction in mortality. However, if they are diverted to those who need them less - perhaps to those for whom the disease has not progressed as far and their risk of opportunistic infection is reduced or to those who have another source and want the security of accumulating a buffer stock - then this will prevent the drugs from having the same national impact on HIV-related mortality.

Studying the impacts of corruption in the context of ARV provision is particularly appropriate for a few different reasons. First, many important outcomes may be only indirectly linked to welfare, whereas the relevant outcome in this study of deaths averted is clearly of direct importance. Second, during this time period there was virtually no domestic production of Antiretroviral Drugs and there is no substitute for these treatments. The next best alternative (good nutrition and treatment and prevention of opportunistic infections through antibiotics) does not have nearly the impact on morbidity and mortality that these drugs do. Therefore, while other studies that look at corruption and goods provision will be unable to measure the entire supply of those goods, this is possible in this case.

Corruption in government could also limit the effectiveness of local supply chains in a number of ways. If promotion within the public sector is not based on performance, there is less incentive for employees to manage transport or work hard at health facilities. Thus the drugs may remain in the country, but sit unused. Similarly, if health facilities are plagued by high absenteeism, drugs may either sit idly or be prescribed with insufficient guidance so that clients are less likely to adhere. Because of its quick rate of mutation HIV is particularly susceptible to the development of drug resistance due to low adherence to the prescribed regimen.[1].

An additional channel through which corruption could influence the impact of imported drugs is through changing allocations within a country. Guaranteeing treatment to those who have low CD4 counts and therefore have the most compromised immune system is the most efficient way to immediately avert deaths using ARVs. However, there is also a benefit to an individual of treatment before the CD4 count is extremely low, and the World Health Organization recently increased the recommended CD4 count threshold of eligibility from 200 to 350. With the higher threshold, demand for the drugs increases and without sufficient supply, other systems of allocation besides targeting those with the lowest CD4 count arise. One notable alternative system of allocation of any health expenditure is political favoritism, including, for example, targeting core supporters or co-ethnics.

[1]It should be noted that adherence to HIV treatment regimens is generally measured to be quite high in developing countries (Mills et al. (2006))

Using data from Kenya about ARV provision before and after an election, I test for politically motivated targeting of new ARV clinics in one country with high corruption levels and high HIV rates. This is done using an original dataset containing all health facilities in Kenya that provide antiretroviral drugs, along with the year in which they began distribution and their GPS locations. This information is linked with ethnicity records to look for evidence of targeting of the placement of ARV clinics in the home area of a newly elected political leader. I find that there are disproportionate clinics opened in areas of the leader's ethnic group. This suggests one mechanism through which corruption reduces the impact of health inputs. Namely, in a country with high corruption, the assignment of ARV clinics follows a political criterion rather than a public health criterion. Further, this pattern of allocation appears to reflect additional clinics added to areas that were already served rather than expanding access to districts that were underserved previously.

This paper is organized as follows. Section 2 discusses the relevant literature and explains how this paper contributes to it. Section 3 outlines the data to be used for the main specification and for the analysis of the case study of Kenya. Section 4 presents the empirical strategy and the results of the cross-country analysis and section 5 discusses the methodology and results of the case study. Section 6 concludes.

2.2 Literature review

Corruption and patronage have long been noticed as an important barrier to effective government provision of services (see for example Mauro (1995) and Bardhan (1997)). Specifically focusing on health, this link between corruption and inefficient provision of services has been used to explain weak observed associations between health spending and health outcomes (Filmer, Hammer and Pritchett (2000), Rajkumar and Swaroop (2008), and Lewis (2006)).

Some researchers have pointed out that one concern with many studies linking health expenditures to outcomes is an insufficient appreciation for displacement of private spending as a result of public investment (Filmer, Hammer and Pritchett (2000), Mauro (1995), and Pritchett (1996)). One driver of this issue is that in many cases, researchers have access to information about what is provided through the public sector, but it is difficult to obtain a complete record of private provision of goods. This is a relative strength of this paper, because the measure of the quantity of antiretroviral drugs comes from information about international imports and should therefore represent the entire flow of antiretroviral drugs into each country to be distributed both through the public and private sectors.

To complement the literature demonstrating links between levels of corruption and disappointing outcomes, there is another strand of literature addressing the potential mechanisms.

Absenteeism and mismanagement are commonly addressed channels (Lewis (2006), Vian (2008)). In a system in which government employment and promotions depend on political patronage or bribes, both the selection and the incentives of public sector workers are limited. A large literature discusses the threats to health outcomes from health worker absence including Banerjee and Duflo (2006), Chaudhury et al. (2006), and Banerjee, Duflo and Glennerster (2008). Specific to antiretroviral drugs, a number of papers have shown the impacts of health-worker absence on long-run impacts of HIV positive individuals, including

through transmission from mothers to children (Goldstein et al. (2013)) and HIV testing (Goldstein et al. (2008)).

Health-worker absence will threaten outcomes if the drugs reach the health facilities where they are to be distributed, but another threat comes from breakdowns in the supply chain. Kangwana et al. (2009) find that malaria medicine is regularly out of stock in health facilities in Kenya because of procurement delays, and Schouten et al. (2011) identify weaknesses in the supply chain of ARVs in Malawi. These could reflect breakdowns in the distribution or centrally.

Interruptions of treatment for HIV from stock-outs or health-worker absence can be extremely costly for the individual and the general population because of the high risk of development of drug-resistance. While this is a threat with the treatment of many infections, HIV is particularly susceptible for three reasons. First, the virus mutates particularly quickly and second, the condition is chronic, so people live for a long time with treatment, presenting a long period of time for the virus to develop drug resistance in each individual. Finally, the development of drug-resistance is particularly devastating for individuals in much of sub-Saharan Africa where third-line and later-line drugs are unavailable so the length of time until drug resistance develops often determines the duration of survival.

An additional channel that exaggerates the role of corruption is the discouragement of public expenditures that comes from a higher effective price of reaching beneficiaries. Besides reducing the impact of government programs, corruption may also increase the cost of provision of goods and services and discourage investment in provision or redistribution (Ferraz, Finan and Moreira (2012), Gupta, Davoodi and Tiongson (2000), Olken and Pande (2012), Olken (2006), Pritchett (1996)). If this is the case, then we would expect that government expenditure decisions would depend on the level of corruption, and indeed Mauro (1998) finds that corrupt politicians induce lower expenditures on education using cross-country variation. On the other hand, there is an alternative strain of literature that argues that corruption may be a more efficient form of taxation that helps an economy overcome initial frictions and grow (see for example Leff (1964) and Huntington and Fukuyama (2006)).

2.2.1 Clientelism and political patronage

Perceptions of corruption may reflect the presence of clientelism. While these two threats to service provision are not the same, they are undeniably linked. Bandiera, Prat and Valletti (2009) distinguishes between *active waste*, characterized by direct gains paid to public officials and corresponding to traditional forms of explicit corruption, and *passive waste*, in which inefficiencies are not due to public officials gaining but other barriers to efficient procurement, including insufficient information or skill. These authors find a correlation between both types of waste, and find that passive waste explains a much larger fraction of government waste in Italy, which may be evidence that the limits on the efficiency of ARV provision in more corrupt countries may not be evidence of bribes being taken or products being syphoned but other factors that inhibit effective procurement and provision. Many types of clientelism may be categorized as *passive*.

There is growing evidence of political patronage in the form of public goods targeted to co-ethnics of leaders. Some of these studies use data from a large number of countries, including Franck and Rainer (2012) who estimate impacts of co-ethnicity with African lead-

ers on primary-school education and infant mortality, and Hodler and Raschky (2010) using satellite images to measure electrification in countries receiving foreign aid. Other studies find this link at the subnational level in South India (Besley, Pande and Rao (Forthcoming) and Besley et al. (2004)). Many of these studies have focused linking outcomes in Kenya with the ethnicity of presidents and ministers. Burgess et al. (2011) find impacts on infrastructure development, although they find that the introduction of multiparty democracy limited the cross-ethnic differences. Other studies find different impacts on a range of outcomes (Posner and Kramon (2011) and Kramon and Posner (2012)). Finally, using a similar methodology Kudamatsu (2009) does not find evidence of co-ethnicity with the leader in Guinea reducing infant mortality. Most of these studies estimate the impacts of co-ethnicities on individual outcomes using large surveys of individuals, although Burgess et al. (2011) measures differences in construction of roads directly. The second empirical analysis in this paper will follow Burgess et al. (2011) and estimate political targeting of health services directly with evidence of placement of new antiretroviral clinics. This type of political targeting, if it is more common in countries with higher levels of corruption could help to explain the link between corruption and inefficient provision of medicine, because targeting based on ethnicity and politics implies not targeting primarily based on need. This will be discussed further below.

2.2.2 Examples from popular press

Many of the theories outlined in the academic literature find anecdotal support in the popular press in articles from a range of African newspapers. These stories address specific examples of leakages of ARVs and funds to support their provision, and other barriers to effective provision of HIV treatment. For example, a report in Zambia found that an enormous fraction of the money that was provided to the country by the Global Fund could not be accounted for.[2] This money could have been used to build clinics and distribution networks to facilitate the distribution of ARVs, but without it, the ARVs would need to be distributed using fewer resources, possibly causing some drugs to go unused, preventing interventions to increase adherence, or hurting the ability of the government to target those who most needed treatment.

A Ugandan newspaper reported that in many areas of Uganda in 2011, ARV clinics had run out of stocks of drugs.[3] If corruption prevents drugs from being restocked in a timely matter, this can have disastrous impacts on HIV mortality, even if the stocks return. First, a lack of consistent adherence to ARVs allows the virus within an individual to develop immunity to the drugs. When treatment is restarted, it is likely to be less effective at preventing opportunistic infections and keeping the individual alive. HIV is known to mutate rapidly, facilitating the development of drug resistance. Second, if this individual is sexually active, this resistance can be transmitted to others. Both factors will reduce the effectiveness of future ARVs, because the lack of consistent supply allows the virus to develop drug resistance.

[2] "Zambia: Corruption scandal rocks ARV programme," PlusNews, Johannesburg, South Africa, March 14, 2011.

[3] Basudde, Elvis (2011), "Where have all the ARVs gone? World Aids Day Supplement" *The New Vision*, Kampala, Uganda, December 2, 2011.

In another article in the same newspaper, alleged corruption prevented a bill to allocate 28.4 billion Ugandan shillings to purchase CD4 count machines. These machines are used to monitor the progress of HIV in an individual and the effectiveness of treatment.[4] With the machines, ARVs could be more efficiently administered. These machines help doctors and clinical officers to determine whether a person has developed drug resistance and ought to be switched to second line treatment, which would improve the effectiveness of the treatment and the likelihood of its prolonging life.

A report from the Ministry of Finance in Swaziland illustrated the scale of money lost due to corruption by showing that it was nearly double the country's yearly budget for social services.[5]

One channel through which corruption could reduce the impact of imported drugs is by preventing targeting based on need in favor of other motives. In order to maximize the benefit of the drugs, they would need to be distributed in such a way that those who can use them have sufficient access that they can begin and successfully adhere to treatment. If corruption allows targeting based on other criteria, this targeting will be weakened, and the drugs may not lead to those most likely to be helped.

In Zimbabwe, PlusNews reports that HIV positive patients are asked to pay providers in order to access drugs which are officially distributed free of charge. [6] With this type of targeting, many who could use the drugs may be denied access in favor of those who either are less likely to be adherent or in less urgent need of the drugs, resulting in higher rates of mortality due to HIV, even with the same quantity of drugs distributed. In order to keep prices high, providers may also restrict access, letting some drugs go unused in order to maximize profits, again resulting in increased mortality.

2.3 Data

2.3.1 Cross-country impacts of corruption and ARVs

The data in this paper comes from many sources. For the first section of the paper, all data is collapsed to a single observation per year in each country in sub-Saharan Africa. The sample is restricted to one region of the world in order to avoid some - but not all - of the standard concerns with cross-country analysis, and to focus on the region that is the hardest hit by the HIV epidemic.

The first datasource is used to measure the quantity of drugs entering each country. This information comes from the WHO Global Price Reporting Mechanism. This is an online database of all international purchases of drugs associated with HIV/AIDS, malaria, and tuberculosis going into developing and middle income countries. For each purchase, the database reports the date of purchase, the country and company of manufacture, the country of the purchaser, and the price and quantity of each type of drug. This contains records for

[4] "Corruption feared in sh28b HIV deal" *The New Vision*, Kampala, Uganda, December 4, 2011.

[5] "SWAZILAND: Corruption exceeds social services budget," IRIN, Mbabane, Swaziland, October 12, 2011.

[6] ZIMBABWE: HIV patients forced to pay up or go without," PlusNews, Harare, Zimbabwe, October 5, 2010.

approximately 30,000 purchases of antiretroviral drugs across three categories - antiretroviral drugs, HIV Diagnostics, and HIV prevention. The same drugs with the same dosages are listed in each category and combining all three reduces threats from misclassification at the international level of how drugs will be used at the local level.[7] The records includes purchases on the part of governments, NGOs, and researchers.[8]

The analysis uses two measures of the quantity of drugs entering each country in each year. The first measure uses standard doses to calculate the quantity of drugs in terms of person-years. Because some drugs are used in combination with others, this measure is imperfect and may be higher in countries that use fewer combination pills. The second measure is the quantity of money spent on all imported ARVs. This is simply the sum of the costs of all purchases.

HIV statistics come from the UNAIDS/WHO 2008 Report on the Global AIDS Epidemic, which for each country in each year reports an estimate of the prevalence, the number of people living with HIV, and the number of deaths due to HIV. While this information may be flawed, there is no better source of information about the prevalence in all countries.

Governance indicators for Control of Corruption, Government Efficacy, and Rule of Law are taken from Kaufmann, Kraay and Mastruzzi (2010).[9] In each year, each country is given a score for each of these indexes. In order to not rely on small differences, the analysis uses binary measures of each of these representing an indicator for a score above the mean in sub-Saharan Africa. GNI per capita is taken from the World Bank's Human Development Indicators.

2.3.2 Case study

The second section of the analysis uses more detailed data obtained from government and private sources in Kenya, combined with population data from MeasureDHS. Information on the placement of ARV clinics is from Kenyapharma, a procurement agency, and the National AIDS and STI Control Program of the Ministry of Health. These reports were provided directly to the author in the Fall of 2011. The information about ethnic backgrounds of populations comes from the 2008/2009 Measure DHS data, in which respondents are asked to report their own ethnicity. The GPS data used to link the two is from the Kenya Open Data Initiative.[10]

[7]For example, antiretroviral drugs purchased as HIV prevention may be used for Prevention of Mother to Child Transmission, and as HIV positive babies typically are overcome by the disease quickly, this should also show up in preventing future deaths.

[8]Agreements between drug companies and developing countries set maximum prices that are low if drugs are purchased by governments or NGOs, but the prices are higher for the private sector. Partly because of this, the private sector does not import large quantities of ARVs in these countries, but it minimally participates in distribution of drugs once they are in the country.

[9]There are many other options of corruption measures, and this measure was chosen because it is both widely used and covers a large number of African countries, which is not the case of many of the more recent measures of actual (rather than perceived) corruption. Other authors have found that the different measures are quite highly correlated, and that the choice of measures rarely changes the outcomes of a study (Svensson (2005), Olken (2009)).

[10](opendata.go.ke)

2.4 Empirical Strategy and Results

2.4.1 Cross-country impacts of corruption and ARVs

This paper will follow previous analysis of cross-country panel-data, including country and year fixed effects and estimating the coefficient on the interaction of corruption and quantities of imported drugs. To ensure some reliability of the data, I will first estimate the following equation to verify that the quantity of drugs is associated with a reduction in deaths due to HIV:

$$deaths_{jt} = \alpha_j + \gamma_t + \beta_1 * ARVs_{jt} + \beta_2 * prev_{jt} + \beta_3 * PLWH_{jt} + e_j$$

where $deaths_jt$ is the number of deaths in year t in country j due to HIV as reported by the WHO, and $ARVs_{jt}$ is the total quantity of drugs imported in that year according to the WHO Global Price Reporting Mechanism. In the first set of specifications, this is included measured in doses (person-years), and in the second set of estimates, the quantity is reported in dollars spent. Controls are included for the prevalence of HIV (α_j)and the number of people living with HIV ($PLWH_{jt}$) as well as country and year fixed effects (α_j and γ_t). The coefficient β_1 shows the association between ARVs entering a country and deaths due to HIV reported in that year.

As reported in columns 1 and 4 of Table 2.1, this is large and negative and statistically significant at all standard levels. Column 1 reports the estimates using the number of person days of drugs as the quantity of ARVs and Column 4 presents the same with the cost of all imported drugs as the measure of quantity. In both cases, the coefficient is negative and significant.

To investigate the role of corruption in changing this effect, I focus on the interaction between the quantity of ARVs and the level of corruption. To do this, I estimate the following equation:

$$\begin{aligned} deaths_{jt} = \alpha_t + \gamma_j + \beta_1 * ARVs_{jt} + \beta_2 * corruption_{jt} \\ +\beta_3 * ARVs * corruption_{jt} + \Sigma b_i * X_{ij} + e_j \end{aligned}$$

where: $ARVs_{jt}$ is the person*years purchased by country j in year t. $Corruption_{jt}$ is an indicator for being below the mean (in Africa) on the Kaufmann, Kraay and Mastruzzi (2010) Control of Corruption Index. To ease interpretation, the quantity of ARVs are demeaned so that the mean is zero. This way the coefficients on the un-interacted terms are meaningful and can be interpreted as the impact at the mean.

If corruption does limit the reduction in deaths generated by purchased drugs, then β_3 should be positive (reflecting a dampened reduction in deaths).

The estimated parameters from this equation are reported in columns 2 and 4 of Table 2.1. In this table, the coefficients on quantities of ARVs are still large and negative and significant, showing that in less corrupt countries, ARVs reduce deaths due to HIV. The coefficient on the interaction term in column 2 is positive, but not significant. A positive

coefficient reflects that the impact of ARVs in corrupt countries is lower, but the fact that it is not significant means that this is inconclusive. In column 4, using spending as the measure of quantity, the interaction term is positive and significant, and large enough to nearly wipe out the impact of ARVs on deaths averted. This suggests that corruption does mitigate the impact of imported ARVs.

Perhaps the variable for corruption is picking up other measures of good governance that have an effect through different channels on the impacts of ARVs on deaths. Columns 3 and 6 of Table 2.1 show the same estimates including the country's GNI and the interaction between that and the quantity of ARVs. With this included, the coefficient on the interaction between ARV quantity and corruption is of a similar magnitude, but significant using both measures of quantity.

Interestingly, the positive coefficients on the interaction between ARV quantity and GNI imply that wealthier countries may see fewer deaths averted as a result of the same quantity of ARVs. One possible explanation is that richer countries have met more of the demand within their countries and the marginal (and average) return is lower as those with less advanced infection are treated. Although not estimated in this paper, these countries may be treating those who would not have died immediately otherwise, but the effects on mortality may show up after a few years.

Table 2.2 further investigates whether the variable for corruption is a proxy for other types of governance. This table adds a number of measures of government quality alone and interacted with ARV quantities. These variables are binary measures constructed in the same way as the measure of corruption coding above average values as 1 and below average as 0.

The first column uses person days of treatment as the quantity and the second uses the price. In column 1, the coefficient on the interaction term of ARVs and Corrupt is still positive and significant, although Good Rule of Law and Effective Governance also have significant relationships when interacted with ARVs. In the second column, only the interaction with corruption is significant.

Are corrupt countries different in other ways? Table 2.3 shows which countries fall in which categories. Table 2.4 compares the countries that are more or less corrupt on a variety of measures. As seen in Table 2.4, more corrupt countries have lower HIV prevalence rates. They also spend more on ARVs, but for a lower quantity. One possibility is that more corrupt countries are buying more expensive drugs that may be less likely to be first or second line treatments, and therefore more valuable since third line and beyond treatments are rarely widely available.[11] Tables 2.5 and 2.6 show the breakdown of types of antiretroviral drugs purchased by more and less corrupt countries. This is measured as the percentage of all drugs purchased in each category these quantities by the specific type of drug.

Based on these comparisons, the quantity of ARVs imported is clearly not exogenous and it is possible that the association shown in this cross-country analysis is not causal. The inclusion of country and year fixed effects deals with many potential threats to endogeneity, but it cannot handle all of them.

[11] ARV treatment becomes ineffective for an individual once the HIV in their system develops resistance to the treatment they are given. Once this happens, a person is given a different treatment, referred to as the second line. In developed countries, this process can repeat many times with those who live with HIV for many years progressing to third, fourth, etc. line treatments.

2.5 Case study in Kenya

If the relationship measured in the previous section is causal, then looking at the mechanisms through which corruption changes effectiveness of imported HIV treatment is a pertinent next step. At the same time, identifying specific mechanisms provides additional support that the relationship is causal.

The fact that controlling for the role of government efficacy does not eliminate the effect is suggestive that the channel through which corruption has an influence is not in simply making government programs less efficient generally with poor incentives for performance or high absenteeism. Instead, this suggests other channels through which drugs are diverted or allocated inefficiently.

One channel through which corruption could reduce the impact of imported drugs is by preventing targeting based on need in favor of other motives. In order to maximize the benefit of the drugs, they would need to be distributed in such a way that those who can use them have sufficient access that they can begin and successfully adhere to treatment. If corruption allows targeting based on other criteria, adequate targeting will be weakened, and the drugs may not reach those most likely to be helped. In this section, I test whether one type of targeting exists in Kenya, a country consistently listed as in the top half of corrupt countries in sub-Saharan Africa.

In particular, I look for evidence of selective placement of ARV clinics in Luo areas after Raila Odinga became Prime Minister in 2008. Previous research has demonstrated that the match between the ethnicity of leaders and constituents is a strong predictor of the provision of public goods (Burgess et al. (2011), Kramon and Posner, 2012). In 2008, after a fiercely contested election for president, followed by allegations of electoral fraud and eventually by violence, the opposition leader, Raila Odinga, became prime minister. Jablonski (2012) looks at government spending in areas populated by Odinga's core supporters after the same election. This paper uses a similar method, focusing exclusively on ARV clinics.

By using a measure of health inputs rather than outcomes, I am able to isolate the impact through expenditures and channeling of resources, and avoid contamination from inspiration. This is similar to Burgess et al. (2011) who look at road construction in Kenya as a function of the ethnic match between the constituents and the government.

If there is targeting based on shared ethnicity, then we would expect to see a relative increase in ARV clinics in Luo areas after the election. To test this, I construct a dataset in which each observation represents one division in one year.[12] For each year between 2004 and 2010 and each of the 225 divisions covered in the 2003 or 2008/2009 DHS survey, this dataset contains the number of clinics which disburse antiretroviral drugs and an estimate of the proportion of the population that self-defines as Luo.

This data is used to look for evidence that Luo areas disproportionately received new clinics after the election, by regressing the number of clinics on the proportion of the population that is Luo, an indicator variable for whether the observation is after the election, and the interaction of the two. I also include controls for the local HIV prevalence at both the district and division levels and year and division fixed effects. The coefficient of interest is the coefficient on the interaction term. Formally, the equation to estimate is:

[12]Kenya has provinces subdivided into districts, further subdivided into divisions.

$$NumClinics_{jdt} = \alpha_t + \gamma_j d + \beta_1 * PercentLuo_{jdt} + \beta_2 * PercentLuo * Post_t$$
$$+\beta_3 Post_t * HIVrate_{jd} + \beta_4 Post_t * HIVrate_d + \epsilon_{jdt}$$

where $NumClinics_{jdt}$ is the number of health facilities distributing ARVs in division j of district d in year t. $Post_t$ is a binary variable that is 1 if the observation is from 2008 or later and 0 if it is earlier. If there exists ethnically-based targeting, one would expect that β_2 would be positive and significant.

Columns 1 and 2 of Table 2.7 shows the estimates of the parameters from the equation above. The coefficient of interest is the coefficient on the interaction between being a year after the election and the percent of the population that is Luo. These are reported in the first row. The first column includes the basic specification without any controls. The second column HIV rates interacted with $Post_t$, the indicator for 2008 and later. In each specification, the coefficient on the interaction term is large, positive and significant. This provides evidence that Luo areas saw a disproportionate increase in the number of HIV clinics after the 2007 election.

To better understand the relationship, I replace the outcome with a binary indicator for whether the division has any ARV clinics, estimating the following equation:

$$I\_Clinic_{jdt} = \alpha_t + \gamma_j d + \beta_1 * PercentLuo_{jdt} + \beta_2 * Post_t$$
$$+\beta_3 * PercentLuo * Post_{jdt} + \beta_4 * HIVrate_{jd} + \beta_5 * HIVrate_d + \epsilon_{jdt}$$

where I_Clinic_{jdt} is an binary variable which takes on a value of 1 in divisions with an ARV-distributing facility in a given year and 0 otherwise.

Columns 3 and 4 of Table 2.7 shows the estimates from this equation. Unlike in the previous table, the coefficients on these interaction terms are consistently insignificant. The estimates are imprecise enough that it is not possible to conclusively rule out some impact on this margin, but the difference between the two tables is suggestive of an increase in intensity rather than an expansion to new areas.

The lack of impact on the extensive margin demonstrated in the last columns of 2.7 is suggestive of a reduction in welfare as a result of this misallocation. Arbitrary distribution of scarce resources may not reduce welfare, but this suggests an increase in distribution without a corresponding increase in access. The degree to which this is true depends on the degree to which the previously existing ARV-distributing facilities were able to meet the local demand.

The response of targeting to the ethnic composition is not likely to be linear as specified above. The analysis from Table 2.7 is repeated, replacing the dependent variable *percent_luo* with an indicator for whether the majority of the population is Luo.[13] The results, reported in Tables 2.8, are qualitatively unchanged.

One explanation for the estimated result is that before the election, Luo areas may have been disproportionately underserved and the increase was bringing them to where they would have been otherwise. Limiting the analysis to the years before the election, Table 2.9 does

[13]The divisions with Luo majorities are in Homa Bay(Kendul Bay, Lake Victoria, Mbita, Ndhiwa, Oyugis, and Rangwe), Kisumu (Lower Nyakach, Muhoroni, Nyando, Upper Nyakach, and Winam), Migori (Migori and Nyatike), and Siaya(Bondo, Boro, Rarieda, Ugunja, Ukwala, and Yala)

not provide evidence that Luo areas were previously underserved.

2.6 Conclusion

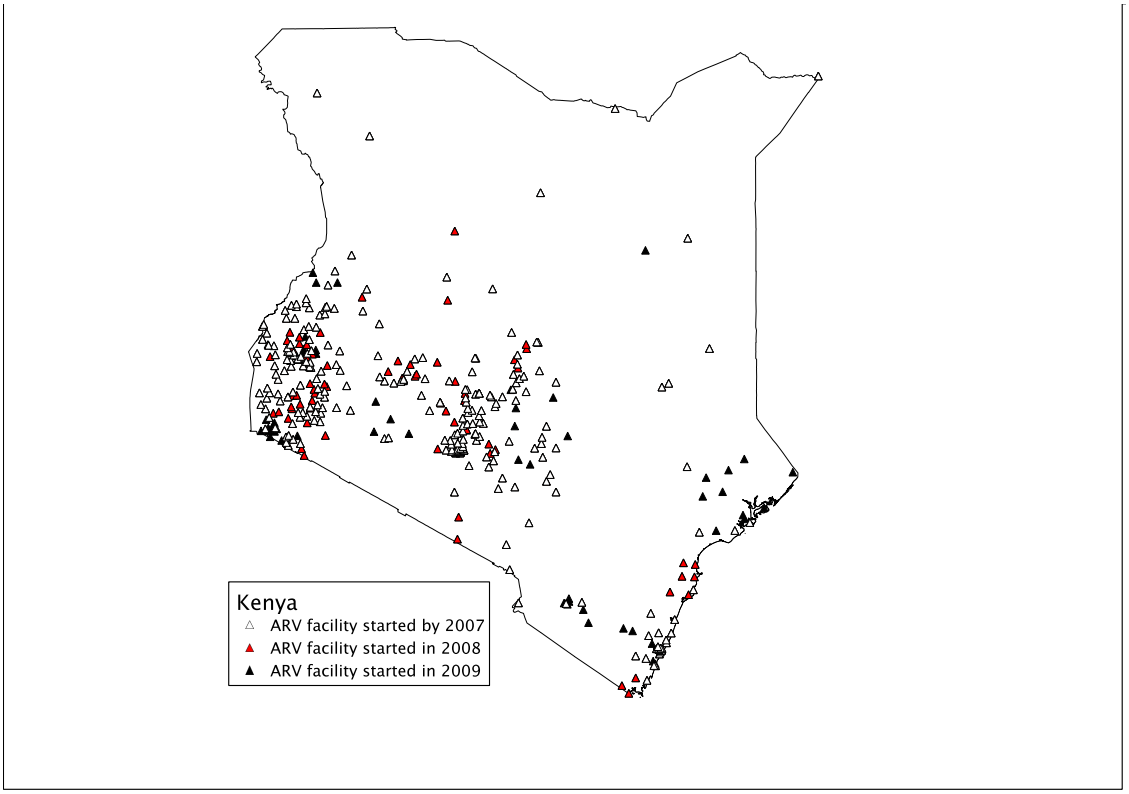
This paper identifies two interesting patterns. First, using a cross-country panel from sub-Saharan Africa, it shows that corruption is associated with a reduction in efficiency of imported antiretroviral drugs. The lack of reduction in mortality from the same expenditure on treatment in relatively more corrupt countries points to a very dangerous consequence of corruption. This fits neatly into a literature on the national impacts of corruption on various outcomes and adds a new outcome with important consequences. Second, using data from within Kenya, I find evidence of political targeting of HIV treatment, again suggestive of a reduction in efficiency of important health expenditures. This speaks to a separate literature on political favoritism and targeting of public goods.

This raises the question of how corruption is maintained. Why are corrupt politicians not voted out of office? Chong et al. (2012) and Banerjee et al. (2012) both find strong evidence that voters punish corrupt politicians, even those of the same ethnicity. Still, targeting resources to constituents may not qualify as corruption and thus it could be an effective strategy to increase electoral support. In addition, while corruption may be punished, Ferraz and Finan (2008) find that low levels of corruption may be tolerated. There are other studies that find evidence of various factors that reduce the level of or the impacts of corruption. For example Reinikka and Svensson (2005) find evidence that an increase in public awareness of leakage of public funds in Uganda reduced capture and improved outcomes, demonstrating that these issues may not be intractable. Similarly Burgess et al. (2011) find that multiparty democracy reduced the degree of ethnic targeting of public expenditures in Kenya. On the other hand, deeply entrenched corruption may be difficult to truly address. Banerjee, Duflo and Glennerster (2008) reports the results of a large-scale government monitoring and incentives program aimed at improving health worker attendance. While the initial results showed improvements, the long-term gains were lost as health facilities learned ways around the rules, and Fisman and Miguel (2007) find that corruption norms follow diplomats, predicting their behavior even outside their home countries.

This paper presents evidence of large costs of corruption in reducing the efficiency of health services that are able to save lives. This is combined with suggestive evidence of one mechanism which may contribute to this relationship. Future work is needed to expand the analysis of mechanisms through which corruption limits the effectiveness of health spending in producing outcomes, and to identify more ways in which these barriers to efficiency can be mitigated.

2.7 Figures

Figure 2.1: ARV clinics in Kenya



2.8 Tables

Table 2.1: Impact of corruption on effectiveness of ARVs

VARIABLES	(1) deaths	(2) deaths	(3) deaths	(4) deaths	(5) deaths	(6) deaths
ARVs (person years)	-0.0821***	-0.0915**	-0.0832**			
	(0.0253)	(0.0392)	(0.0409)			
ARVs*Corrupt		0.0356	0.0346*			
		(0.0254)	(0.0202)			
ARVs*High GNI			0.168***			
			(0.0440)			
Spending on ARVs (1000s)				-0.384***	-0.716***	-0.601***
				(0.140)	(0.0876)	(0.108)
Spending*Corrupt					0.434***	0.396***
					(0.120)	(0.107)
Spending*High GNI						1.456*
						(0.834)
Corrupt		4,318**	3,538**		4,705***	3,626***
		(1,825)	(1,487)		(1,173)	(1,189)
High GNI			4,629***			6,249***
			(1,414)			(1,574)
HIV prevalence	1,076	1,113	2,718**	941.1	636.5	2,416**
	(1,805)	(1,875)	(1,291)	(1,710)	(1,501)	(946.5)
PLWH	0.0766**	0.0801***	0.0540***	0.0653**	0.0807***	0.0555***
	(0.0298)	(0.0295)	(0.0170)	(0.0259)	(0.0220)	(0.0113)
Observations	142	142	142	142	142	142
R-squared	0.998	0.998	0.998	0.998	0.998	0.999

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
All estimates include both country and year FEs,
and all SEs are clustered at the country level.

Table 2.2: Impact of corruption and other government indicators on effectiveness of ARVs

VARIABLES	(1) deaths	(2) deaths
ARVs (person years)	-0.0621	
	(0.0982)	
Spending on ARVs (1000s)		-0.656***
		(0.225)
ARVs*Corrupt	0.0148	
	(0.0918)	
Spending*Corrupt		0.410***
		(0.128)
ARVs*High GNI	0.177***	
	(0.0504)	
Spending*High GNI		1.427
		(0.855)
ARVs*Good Rule of Law	0.0396	
	(0.168)	
Spending*Good Rule of Law		-0.0832
		(0.329)
ARVs*Effective Governance	-0.0593	
	(0.112)	
Spending*Effective Governance		0.136
		(0.302)
Corrupt	3,064	3,452**
	(2,389)	(1,472)
High GNI	4,659***	6,588***
	(1,259)	(1,411)
Good Rule of Law	376.8	-664.0
	(3,633)	(2,368)
Effective Governance	-1,730	-760.8
	(2,568)	(1,636)
HIV prevalence	2,756**	2,172**
	(1,341)	(1,032)
PLWH	0.0526***	0.0571***
	(0.0175)	(0.0118)
Observations	142	142
R-squared	0.998	0.999

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
All estimates include both country and year FEs,
and all SEs are clustered at the country level

Table 2.3: Countries by corruption status

	Never Corrupt	Sometimes Corrupt	Always Corrupt
Never Rich	Burkina Faso	Benin	Burundi
	Eritrea	Comoros	Cameroon
	Ghana	Djibouti	Central African Republic
	Lesotho	Ethiopia	Chad
	Madagascar	Gambia	Cote d'Ivoire
	Mali	Guinea	Guinea-Bissau
	Mauritania	Liberia	Kenya
	Mozambique	Malawi	Niger
		Rwanda	Nigeria
		Togo	Sierra Leone
		Zambia	Sudan
			Uganda
Sometimes Rich	Senegal	Sao Tome and Principe	Angola
			Zimbabwe
Always Rich	Botswana	Gabon	Republic of the Congo
	Cape Verde	Tanzania	Democratic Republic of the Congo
	Namibia	Equatorial Guinea	
	Seychelles	Somalia	
	South Africa		
	Swaziland		

Table 2.4: Comparison of more and less corrupt countries

Variable	Less Corrupt	More Corrupt
HIV prevalence	7.442	4.456
	(8.768)	(4.467)
PLWH	607633.33	444457.65
	(1320124.3)	(624440.62)
ARVs (person years)	26416.174	24679.665
	(60638.485)	(60317.787)
Money for ARVs (1000s)	5481.701	7041.231
	(11247.696)	(15658.17)

Table 2.5: Drugs purchased by country, by dose

Variable	Less Corrupt	More Corrupt
Abacavir (ABC)	1.67	.55
Atazanavir (ATV)	0	.
Combination	32.59	19.44
Didanosine (ddI)	3.91	2.93
Efavirenz (EFV)	24.46	25.28
Indinavir (IDV)	0	.
Lamivudine (3TC)	10.3	1.33
Nelfinavir (NFV)	.18	2.49
Nevirapine (NVP)	14.74	38.63
Ritonavir (RTV)	.1	1.33
Saquinavir (SQV)	.04	0
Stavudine (d4T)	3.19	5.6
Zidovudine (ZDV)	8.82	2.43

Table 2.6: Drugs purchased by country, by money spent

Variable	Less Corrupt	More Corrupt
Abacavir (ABC)	7.09	1.38
Atazanavir (ATV)	1.45	.
Combination	42.55	44.15
Didanosine (ddI)	5.41	2.96
Efavirenz (EFV)	25.17	29.05
Indinavir (IDV)	.53	.
Lamivudine (3TC)	3.64	.36
Nelfinavir (NFV)	1.2	10.15
Nevirapine (NVP)	4.45	8.6
Ritonavir (RTV)	.19	.46
Saquinavir (SQV)	.56	1.36
Stavudine (d4T)	1.25	1.16
Zidovudine (ZDV)	6.19	.36

Table 2.7: Targeting of introduction of ARVs in health facilities in Kenya

VARIABLES	(1) Num. ARV clinics	(2) Num. ARV clinics	(3) Any ARV clinics	(4) Any ARV clinics
Post*PercLuo	2.079***	1.753**	0.0442	-0.0681
	(0.528)	(0.874)	(0.0760)	(0.161)
Post*HIVdivision		1.553		-0.132
		(2.565)		(0.306)
Post*HIVdistrict		0.727		0.774
		(4.841)		(0.685)
Observations	1,568	1,260	1,568	1,260
R-squared	0.721	0.724	0.747	0.748
Clusters	224	180	224	180

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Standard errors clustered at the division level. All estimates include division and year FEs.

Table 2.8: Targeting of introduction of ARVs in health facilities in Kenya

VARIABLES	(1) Num. ARV clinics	(2) Num. ARV clinics	(3) Any ARV clinics	(4) Any ARV clinics
Post*LuoMajority	1.906***	1.631**	0.0540	-0.0388
	(0.463)	(0.714)	(0.0705)	(0.131)
Post*HIVdivision		1.635		-0.144
		(2.588)		(0.296)
Post*HIVdistrict		0.792		0.684
		(4.496)		(0.638)
Observations	1,568	1,260	1,568	1,260
R-squared	0.721	0.725	0.747	0.748
Clusters	224	180	224	180

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Standard errors clustered at the division level. All estimates include division and year FEs.

Table 2.9: Previously underserved?

VARIABLES	(1) Num. ARV Clinics	(2) Any ARV clinics	(3) Num. ARV clinics	(4) Any ARV clinics
LuoMajority	0.237	-0.0392		
	(0.380)	(0.163)		
HIVdivision	0.130	-0.0992	0.0423	-0.125
	(0.858)	(0.383)	(0.829)	(0.374)
HIVdistrict	2.501	1.148	1.847	0.914
	(1.630)	(0.849)	(1.617)	(0.896)
PercLuo			0.448	0.0288
			(0.404)	(0.193)
Observations	720	720	720	720
R-squared	0.212	0.266	0.215	0.266

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Standard errors clustered at the division level.
All estimates include division and year FEs.

Chapter 3

Local Economic Conditions and Participation in the Rwandan Genocide

3.1 Introduction

The one hundred days in 1994 in which nearly one million Rwandans were brutally killed by their neighbors is widely considered to be one of the most horrific moments of the last century. The enormous number of participants is baffling. This paper contributes to a small but growing literature exploiting subnational variation in participation in conflict in an attempt to make comprehensible such astonishing events.

This paper addresses three sets of theories about participation in violence. These categories, discussed below should be considered *ideal types* acknowledging that the distinctions between them are contrived. While many theories of violence incorporate elements of multiple theories, the distinctions provide a useful framework for discussion. Each of the categories below is defined by the empirical predictions of the theories it contains.

The first category will be referred to as the *Opportunity Cost* theories. These theories focus on the low expected value of alternative endeavors on the part of potential participants in violence. These theories have a strong tradition in many social sciences, from economics and political science focusing on explicit opportunity costs of engaging in destructive activities, to sociology, focusing on a weakened desire to continue to participate in society with rules that may not be perceived to provide any benefit to the individual. These theories predict that the lower the level of resources held by potential aggressors or the fewer the opportunities he or she has to improve this or her welfare through productive endeavors, the more likely he or she will be to participate in violence. If these theories are true, then increases in employment and education should all be associated with a reduction in participation in violence.

The next set of theories will be referred to as the *Relative Wealth and Expropriation* theories. These theories, which have often been combined with the first category, focus on characteristics of those towards whom violence is directed. When potential victims of violence have more, this can generate dissatisfaction with the situation or simply increase

the payoff to expropriating their resources. Such theories would predict that if the target of potential violence has either more opportunities that can generate jealousy or greater assets that can be appropriated, there is likely to be more violence observed.

The final set of theories will be referred to as the *Political Participation and Unmet Expectations* theories, which propose to explain observed associations between education and violence. These theories conceive of violence as a form of political participation, argue that education increases expectations of economic success that generate frustration when unmet, or claim that those who are more educated can more easily be mobilized for any type of mass action, including violence. Such theories have often been proposed in response to observed patterns of participants in violence that were inconsistent with earlier theories, most notably: support for the violent elements of the Nazi party during the Holocaust, participation in urban unrest in the US in the 1960s, and recently, participation in terrorism. These theories predict that those who are most likely to participate in violence may not be those with the fewest alternatives, but instead the most educated.

The Rwandan genocide presents both an enormous argument in support of the need to study causes of violence and a potential opportunity to do so. This paper tests the previously mentioned theories using a dataset that combines recently released data on local participation in the genocide and census data collected before the start of the genocide. The data on participation comes from records of the number of individuals accused of crimes associated with genocide, kept by Gacaca tribunals, constructed to prosecute the enormous number of individuals who were still awaiting trial ten years after the genocide. This data is combined, at the local level, with the 1991 census, which contains information on education, employment, a range of demographic characteristics, and ethnic identification.

Using this combined data, I find that participation is associated with greater education among Hutu, and lower employment among Hutu, but is not significantly associated with characteristics of the education and employment status of the Tutsi population. This is robust to the inclusion of a large set of controls, including age and migration patterns of the population, radios, and province fixed effects.

The paper proceeds as follows: The next section lays out the theoretical framework, outlining predictions of the theories to be empirically tested, and situates this paper within the broader literature on causes of violence and civil conflict. Section 3 briefly describes the historical context as is relevant to this project. Section 4 outlines the data to be used and section 5 specifies the empirical strategies. Results are presented and discussed in section 6, and section 7 concludes.

3.2 Theoretical Framework and Literature Review

This section will outline each of the three sets of theories of causes of violence that will later be tested. Each subsection will contain a discussion of the theories, their support and challenges in earlier research, and the predictions the theories would make about who is most likely to participate in violence. The final subsection will discuss related empirical research on violence.

3.2.1 Opportunity Cost Theories

The opportunity cost theories posit that those with the least to lose by doing so are the most likely to participate in violence. It is assumed that violence is anti-social and its enactment crowds out productive activities. Thus, holding other factors constant, an individual with a low pay-off to productive work - because of low ability or limited opportunities, for example - is more likely to engage in violence.

This idea is key to models of the individual choice between production and appropriation (Haavelmo (1954), Grossman and Kim (1995), Skaperdas (1992), and Hirshleifer (1995)) and models of crime beginning with Becker's seminal work (Becker (1968)). In such models, individuals who expect little benefit from production are more likely to expropriate the resources of others or to engage in crime as a means of obtaining the resources that could not otherwise be obtained. A related theory with similar predictions arose in sociology, explaining individual willingness to participate in violence as a step away from a society whose rules do not benefit that individual. Merton (1934) is often attributed with presenting anomie, and the inability to construct well-being through legitimate means, as a cause of deviance.

More recent work has extended rational choice models to explain rebellions, empirically testing whether rebellion could be explained as a response to labor market constraints. Such theories propose that increases in income, education, and economic growth - which increase the opportunity cost of fighting - should reduce the likelihood of rebellion. Collier and Hoeffler (2002) find that at the cross-country level, higher income and higher levels of education are associated with a lower risk of rebellion through an increase in the opportunity cost of violence. They argue that in a society with more education and income, political change can happen through institutional pathways. Similarly, Fearon and Laitin (2003) find an association between weak states and political violence, arguing that weak states provide little opportunity to enact political change through non-violent channels, which can also reduce the opportunity cost of violence.

A large number of cross-country studies have found a connection between poor economic conditions and civil conflict, using a range of identification strategies (Collier and Hoeffler (2004), Fearon and Laitin (2003), Hegre and Sambanis (2006)). These studies rely on national income data and are generally unable to distinguish between variation in resources held by aggressors and victims. Miguel, Satyanath and Sergenti (2004) use rainfall as an instrument for economic growth, which presumably works through agriculture, a likely determinant of the opportunity cost of fighting among those most likely to fight.

Recent work that uses data on subnational variation in violence has supported this hypothesis that negative shocks to income among potential aggressors facilitate violence (Justino (2009)). Do and Iyer (2010) look at local correlates of violent conflict in Nepal and find that conflict intensity is associated with higher levels of poverty. Murshed and Gates (2005) also look at local-level predictors of violence intensity in Nepal and find that areas with greater landlessness and lower HDI indicators experienced greater levels of violence. Humphreys and Weinstein (2008) find that rebel organizations in Sierra Leone that rely on voluntary recruitment, individuals with limited economic options outside of joining rebel groups are most likely to join.

The theory that opportunity costs reduce willingness to participate in violence rests on the

assumption that individuals cannot both work and rebel, and that what they have without rebelling is put at risk if they choose to rebel. For this reason, the time spent in employment may have a stronger deterrent effect than do the wages earned. Dube and Vargas (2013) use variation in international coffee prices and international oil prices to identify different effects of demand shocks in areas in which a labor-intensive industry (coffee) is central and areas dominated by an industry that does not rely on large labor investments (oil). They find that negative shocks to international coffee prices lead to more violence in coffee-producing areas - which is consistent with the opportunity cost theories - and that positive shocks to oil prices lead to more violence in oil-producing regions, which is consistent with the suggestion that resources to finance violence can increase violence, as well as with theories to be discussed below.

Fighting can also be costly for an individual who is not employed but anticipates the possibility of becoming employed by removing this option. Thus, opportunity costs can change the likelihood of participation even for those who are not at the moment employed, if they live in an area with functioning labor markets in which they plan or hope to participate, and to which they believe they will be able to gain access.

While such opportunity cost theories may appear broadly plausible, empirical tests are threatened because motivation for violence is not the only channel through which the resources of potential attackers can influence the likelihood of violence. In particular, even if more resources may decrease the motivation for violence, more resources, even those held by potential attackers, may also make violence more feasible. Mitra and Ray (2010) design a model in which greater resources on the part of the attacker both reduce the desire for conflict and increase the feasibility of conflict. The prediction of their model is thus inconclusive regarding the sign of the effect of resources held by potential attackers on the likelihood of attacks. Collier and Hoeffler (2004) find that the availability of resources to finance conflict is associated with greater levels of violence. Berman et al. (2011) find evidence that is inconsistent with the opportunity cost theory of violence, showing that in their data, unemployment in Iraq and the Philippines is not associated with an increase in attacks on government or attacks on civilians. However,another paper also using data from Iraq finds a negative relationship between quasi-random variation in employment availability in Iraq and insurgency frequency (Iyengar, Monten and Hanson (2011)).

In the case of Rwanda, both coffee and tin prices fell dramatically just a few years before the onslaught of violence, and a rapidly growing population increased demand for already limited land for farming. The claim that Rwanda was the victim of a Malthusian crisis in which limited resources made the growing population unsustainable and unstable (Boudreaux (2009), Diamond (2005), Yanagizawa (2006)). Soil fertility had "fallen sharply" due to over-cultivation (Percival and Homer-Dixon (1998)). One author writes: "Of the nearly 60 percent of Rwandans under the age of twenty, tens of thousands had little hope of obtaining the land needed to establish their own households or the jobs necessary to provide for a family." (Des Forges et al. (1999)) These factors may have contributed to a reduction in the opportunity cost of violence that facilitated the high rates of participation observed: "Many of these zealous killers were poor, drawn from a population 86 percent of whom lived in poverty, the highest percentage in the world. They included many young men who had hung out on the streets of Kigali or smaller commercial centers, with little prospect of obtaining either the land or the jobs needed to marry and raise families." (Des Forges et al.

(1999)). Des Forges et al. (1999) describes the lure of fighting for those with little otherwise: "Authorities offered tangible incentives to participants. They delivered food, drink, and other intoxicants, parts of military uniforms and small payments in cash to hungry, jobless young men. They encouraged cultivators to pillage farm animals, crops, and such building materials as doors, windows and roofs. Even more important in this land-hungry society, they promised cultivators the fields left vacant by Tutsi victims. To entrepreneurs and members of the local elite, they granted houses, vehicles, control of a small business, or such rare goods as television sets or computers."

To summarize, the opportunity cost theories posit that resources and opportunities to obtain resources through productive endeavors limit the motivation for violence. If these are true, we will expect to see higher levels of education and employment on the part of potential attackers associated with lower rates of participation in the genocide.

3.2.2 Expropriation and Relative Deprivation theories

The next category of theories deal with the resources held by individuals who are potential victims of violence. These theories suggest that when the potential victim has more resources the likelihood of violence is higher because of either greater dissatisfaction with relative differences in wealth or opportunity, or more directly through a higher expected value of expropriated resources. Such theories are often discussed in combination with ideas from the first set of theories under the umbrella of inequality or relative deprivation theories.

The motivation for expropriation presented by more valuable resources is another key element of classic "guns or butter" models of conflict as a choice between production and appropriation (Haavelmo (1954), Becker (1968), Grossman and Kim (1995), Skaperdas (1992), Hirshleifer (1995), Collier and Hoeffler (2004), Dal Bó and Dal Bó (2011)). In recent work, Mitra and Ray (2010) model inter-group conflict in which one group is predetermined as the aggressor and the other as the victim. Their model predicts that the resources of the victim should unambiguously increase the likelihood of violence between the two groups as violence is at least partially assumed to be undertaken for the purpose of extracting resources from the victim's group. These theories predict that for purely material reasons, resources held by potential victims increase the likelihood of their being attacked.

Others have focused on the role of the resources held by potential victims in generating dislike among those who have less, even when these resources cannot be appropriated. Relative deprivation theories and theories of the relationship between inequality and conflict argue that when potential victims have more than potential attackers, an increase in the resources held by potential victims can generate violence. Gurr (1970) argues that disappointing comparisons with others in the same society can inspire rebellion. Scott (1977) explains agrarian rebellions through dissatisfaction with growing inequality that did not meet norms of agricultural societies, and Sen (1973) points to inequality as a key driver of rebellions. More recently, Chua (2004) studies the role of unequal economic and political power held by "market dominant minorities" and resentment of relatively less well-off majorities in fueling ethnic conflict.

There are many studies empirically linking inequality with increased participation in violence, and inequality between the two ethnic groups has been a key element of many narratives of the Rwandan genocide. Many papers have focused on the relationship between

income inequality and crime or other types of violence within US cities (blau1982cost, Kelly (2000)), within cities elsewhere (Demombynes and Özler (2005), Østby, Nordås and Rød (2009)), and between countries (Alesina and Perotti (1996), Fajnzlber, Lederman and Loayza (2002)). Inequality has also been linked to civil war and political violence in part through an increase in the material gain of expropriation (Blattman and Miguel (2010), Dube and Vargas (2013), Murshed and Gates (2005)).

On the other hand, another set of studies has found no relationship between overall inequality and conflict (Lichbach (1989), Collier and Hoeffler (2004), and Fearon and Laitin (2003)) and others have argued that the studies that did find significant results rely on weak identification and suffer from problems of endogeneity (Land, McCall and Cohen (1990), Soares (2004)).

Still others have argued that - rather than overall inequality - between-group or horizontal inequality may be a more important predictor of civil conflict (Sen (1973), Tilly (1999)), and some have recently found empirical support for this relationship globally (Østby (2008), Cederman, Wimmer and Min (2010), Cederman, Wimmer and Min (2010)). Macours (2011) used subnational data in Nepal, and found that recruiting via youth abductions were higher in areas with increasing land inequality, arguing that relative deprivation, and not only absolute deprivation contributed to the level of conflict.

In the specific case of the Rwandan genocide, inequality has played a prominent role in many discussions. The story of the poor and downtrodden Hutu peasants overthrowing their royal Tutsi oppressors has been a standard part of the Rwandan story, or what some would call a myth, since long before the 1959 revolution in which Hutu took control of the government away from the long-standing Tutsi rulers, and it has continued to come up as a key part of the story of the genocide (Mamdani (2002), Prunier (1995), Straus (2006)). The popular press generally mentions the Belgian elevation of the Tutsi as superior and the subsequent belittling of the Hutu when discussing the genocide. Specific accounts rarely fail to mention both the possibility of looting and the existence of jealousy between groups to explain what happened in 1994 (Cramer (2003), Des Forges et al. (1999), Gourevitch (1998), Hatzfeld (2005), Hatzfeld (2013), Mamdani (2002), Prunier (1995), Straus (2006)). Still, many caution that the purely material motivation of taking Tutsi property was likely to have been minimal (Des Forges et al. (1999), Prunier (1995), Straus (2006)).

Regarding jealousy, Prunier (1995) writes: "Political power had been in Hutu hands for thirty-five years but, thanks to the Belgian social and educational favouritism towards the Tutsi for the forty years before that, the Tutsi community was still able to do well for itself socially and economically. This did not only mean the big Tutsi businessmen; it also meant that most of the local personnel in foreign embassies and in NGOs and international agencies were Tutsi, that there were many Tutsi in the professions and even that the best and highest-priced bar girls, the ones to be encountered in the big hotels, were Tutsi. Social envy came together with political hatred to fire the Interahamwe bloodlust."

Regarding the possibility of looting, "There was of course also an element of material interest in the killings, even in the countryside. The killers looted household belongings and slaughtered the cattle. Meat became very cheap, and grand feasts were held, as if in celebration of the massacre. ... Villagers also probably had a vague hope that if things settled down after the massacres they could obtain pieces of land belonging to the victims, a strong lure in such a land-starved country as Rwanda." (Prunier (1995)). Des Forges et al.

(1999) corroborates the description of the incentive of what could be obtained: "The killing campaign created new opportunities for getting rich as Tutsi property became available for appropriation and it generated new possibilities for acquiring power as political alliances shifted." (Des Forges et al. (1999)).

This class of theories predict that, holding other factors constant, an increase in the resources held by potential victims is likely to be associated with higher participation rates.

3.2.3 Political Participation and Unmet Expectations

The final set of theories conceive of violence as a form of political action and acknowledge that education can generate dissatisfaction. In response to dissatisfaction, individuals may respond violently in order to change the distribution of power. These theories do not dispute the above theories, but rather propose additional - and sometimes complementary - ways in which the distribution of resources between individuals can generate conflict, sometimes leading to different empirical predictions. Theories of violence as political participation typically rely on the presence of dissatisfaction and it is important to note that this dissatisfaction can be generated in part by poverty and relative deprivation. In many cases, these theories of violence as a form of political participation arose in order to explain characteristics of individuals or groups undertaking violence that did not fit the predictions of the previous models.

A common theme of stories of violence as political participation is that individuals who have reason to hope for good economic outcomes - either through investments in education or through rapid political, social, or economic advancement - are confronted with disappointing available options - high unemployment, high poverty, or continued discrimination or political repression. One early example of this was the case of urban civil unrest in the United States during the 1960s. Previous theories had predicted that those involved in civil unrest would be the least well-off and the most-marginalized. Yet high levels of participation among well-educated and well-connected individuals called these theories into question, spawning a broad literature in sociology and political science in which researchers attempted to produce new theories to explain the surprising make-up of the participants (Mason and Murtagh (1985), Sears and McConahay (1973)). The J-Curve theory (Davies (1974), Miller, Bolce and Halligan (1977)) was a popular version of the theory that unmet expectations contributed to increased participation in violence among those who were not least well off. The theory proposed that rapid progress followed by a decline leads people to revolt. Progress raised expectations and the decline crushed them. Gurr (1970) proposed similar explanations for rebellions more generally, arguing that unmet expectations could generate a feeling of disappointment relative to what was anticipated, leading to violence. In these conceptions, when realized welfare outcomes are more beneficial for individuals, this may reduce conflict, but when expected outcomes are higher, this could increase conflict. In particular, more education may increase expectations, and - in the absence of available employment - may lead to violence.

More recently, similar theories have been generated to explain a related and also surprising observation made about terrorism: those who belong to terrorist organizations and even those who participate in suicide bombings are often not the least educated or most marginalized. Krueger and Maleckova (2003) find that poverty and low education do not predict individual

participation in terrorism. Berrebi (2007) finds that high education and income are both positively associated with participation in terrorist organization. Abadie (2006) finds that poverty does not predict cross-country levels of terrorism, but that political dissatisfaction does matter. These all present strong evidence of a stylized fact: although most models of violence or crime predict that all forms of marginalization will lead to increased participation in violence, there are many instances in which those who participate are far from the least well-off. Kopstein and Wittenberg (2011) find evidence of an association between high levels of education and support for the Nazi party in Eastern Europe before World War II.

These correlations between increased education and participation in terrorist groups are threatened by the presence of selection. If terrorist groups only allow those who are more capable to join, then the observed patterns of membership could reflect choices made by the organization rather than the individual. Recent work has focused on disentangling selection and motivation (De Mesquita (2005)). In the case of the Rwandan genocide, selection is not an issue because there was very little or no selection: radios broadcast invitations to all to join and the tools used were those available to farmers, rather than scarce or difficult to use modern weapons.

Another way that education could generate an increase in participation in violence is if education facilitates organization, which could have facilitated larger groups of people participating in an ongoing effort. With this theory, it is not necessary that education increases tensions, but it simply increases the likelihood of collective action being taken.

In Rwanda, education - which had once been mostly a privilege of the ruling Tutsi - had been increasing dramatically among Hutu since Hutu leadership had taken control of the government in 1959 (Prunier (1995)). Economic growth was also relatively high in the 1970s, "but the prosperity was both fragile and superficial. The mass of the people stayed poor and faced the prospect of getting only poorer." (Des Forges et al. (1999)). These initial advances may have increased expectations which were disappointed.

This set of theories propose that higher levels of education among potential aggressors may in some cases be associated with greater levels of violence.

3.2.4 Literature on Violence and Civil Unrest in Rwanda

This project can also be situated within a small literature in economics and political science explaining participation in the Rwandan genocide. The empirical literature on the economics of violent conflict in the world is notably lacking in micro-studies to understand individual motivations for participation in violence (Blattman and Miguel (2010)), and Rwanda is no exception.

Verwimp (2004) used a database of survivors and victims from a single prefecture (Kibuye) to identify individual predictors of survival, including age, sex, occupation, and actions taken after the start of the genocide. Justino and Verwimp (2008) also find high excess mortality among the wealthiest regions in Rwanda. In a separate paper, Verwimp (2005) constructs a panel of data for 350 households interviewed before and after the genocide and finds that individuals with higher off-farm income and those from households with higher incomes were more likely to participate in the genocide. This is partially in contrast with the conclusions of André and Platteau (1998) who focus on characteristics of victims. They find that those with greater land-holdings, as well as those who are considered trouble-makers, are more

likely to be victims, and they conclude that the genocide was a time for the settling of scores and the undertaking of land grabs. The different findings of the two studies may be partly explained by the different samples examined, the latter using a database that is nearly exclusively Hutu. Verwimp (2005) also does find that the percent of land that a household rents is positively associated with the likelihood of participtionn. Looking at 1993 massacres in Burundi, Bundervoet (2009) finds that those with more livestock and other resources were more likely to be attacked. De Walque and Verwimp (2010) use the 2000 Demographic and Health Survey to estimate excess mortality and thus correlates with the likelihood of dying during the genocide. They find that those in urban areas and with more educated backgrounds were more likely to be killed. They do not look at characteristics of the killers but contributes to the story about motivations for killing. Yanagizawa-Drott (2010) recently took advantage of the newly released Gacaca data (described below) and the presence of mountains that can block radio signals to look at the impact of exposure to hate radio on participation rates and found large impacts. This followed an earlier paper that used more limited data and found a smaller - almost negligible - effect of exposure to radio on participation (Straus (2006)).[1]

3.3 Historical Context

Many key facts about the genocide are debated, from the number of people killed to who shot down the president's plane that set off the start of the massacres. Here I will attempt to very briefly summarize the key events that are necessary to provide background for this study. More detailed accounts can be found in Des Forges et al. (1999), Gourevitch (1998), Mamdani (2002), and Prunier (1995).

At the time of independence from Belgium, the Rwandan population was almost entirely comprised of two ethnic groups, the Tutsi and the Hutu. The nature of the distinction between these two groups before colonization is heatedly debated. Early accounts argued that pastoralist Tutsi were the descendents of Hamitic migrants coming from either Egypt or Ethiopia and that the Hutu were agriculturalist Bantu who had lived in the area for much longer. Others have since argued that the physical differences in height and facial features between the two groups noticed by the colonizers were simply the result of different diets one rich in milk and the other relying exclusively on agricultural produce and that the two groups shared a common ancestry. There seems to be universal agreement that the Belgian colonizers at least accentuated and hardened these differences by identifying the Tutsi as a superior race of outsiders born to rule over the Hutu peasants. Although some have argued that the differences between the ethnic groups were small before the genocide, despite having a common language (Kinyarwanda), a common religion (Christianity), and common neighborhoods (the proportion of Tutsi in a commune is never more than forty percent), the intermarriage rate was low before the genocide 28.6 percent of Tutsi were

[1]There are many extremely important debates about causes of civil conflict and the nature of the genocide in Rwanda that this paper does not address. For example, this paper will not touch on debates about whether violence fits the pattern of a genocide or of politicide. It will take ethnic identity as pre-determined and reasonably stable. Details of recruitment methods and internal organization of fighters or the psychological factors that allow for mass violence to occur will not be addressed. It will provide empirical tests of the predictions of three models of the relationship between resources and participation in civil conflict using data from the Rwandan genocide of 1994.

married to Hutu and 2.5 percent of Hutu were married to Tutsi and it had not increased in a generation. Of those in the census, 26.8 percent of Tutsi parents were married to Hutu and 2.4 percent of Hutu parents were married to Tutsi. [2]

In 1959, Tutsi leadership was overthrown in a "democratic" revolution that installed a Hutu government. After this and later small massacres, many Tutsi began living in exile in Burundi, Uganda, and Zaire/the Democratic Republic of the Congo. Some of these exiles formed the Rwandan Patriotic Front in Uganda in 1987, which agitated, along with Tutsi within Rwanda, for representation in the government. At the same time, many other opposition parties within Rwanda began to fracture with some branches supporting "Hutu Power." In 1993 an agreement in Arusha was reached to incorporate opposition parties and the RPF into the Rwandan government, but the existing government, led by President Juvenal Habyarimana, was slow to implement the accords.

On April 6, 1994, Juvenal Habyarimana boarded a plane to Kigali with the President of Burundi, coming from a meeting with other East African leaders to encourage the President to implement the protocols from the Arusha accords which outlined the construction of a new government that would incorporate the RPF leadership and representatives from other opposition political parties. Immediately before landing in Kigali, the plane was shot down and landed in the president's compound.

That evening, killings of Tutsi and moderate Hutu leaders began and massacres of Tutsi and moderate Hutu quickly began across the country, lasting for 100 days. Estimates of the number killed range from 500,000 to more than 1 million.[3] Estimates of the number of participants in the violence are similar or even larger, with the majority of the killings having been committed with machetes and clubs.

Those accused of the highest levels of crimes - organization of massacres - are being tried in the International Criminal Tribunal for Rwanda in Arusha, Tanzania. However, those accused of lower level crimes - including killing and smaller-scale organizing - are tried in local courts, known as Gacaca tribunals, which are based on a traditional system of justice in Rwanda. These courts are set up at the sector level and involve those who are prosecuted confessing their crimes to the community. These tribunals keep very detailed records, and it is the numbers of accused in each sector that provide the basis for this analysis.

3.4 Data

The data is taken primarily from two sources. First, participation rates come from records produced by local-level tribunals, or Gacaca courts. These records give the number of individuals accused of committing each of three levels of crimes during the genocide in each sector (village). Three other primary sources of data at different geographical levels or for smaller numbers of sectors are used to test the validity of this data. Second, pre-genocide education and employment information is taken from the 1991 census in Rwanda, available through IPUMS International.

[2]The difference between the intermarriage rates for those of each group comes from the different sizes of the populations of each ethnic group within the country.

[3]The Government of Rwanda generally uses 1.2 million and most major international media outlets (including Al Jazeera, BBC, New York Times, Time Magazine, the Times (of South Africa)) use 800,000.

3.4.1 Participation Rates

I will use the number of people accused by Gacaca courts in a sector as a proxy for the number of people who participated in the genocide in that sector. These decentralized courts were established, based on a traditional system of participative justice, in order to expedite the process of trying the enormous number of people accused of committing acts of genocide who were still awaiting trial years later. In 2005, at the local level, detailed lists were constructed of those to be prosecuted. The accused can be prosecuted for one of three categories[4] defined by the Government of Rwanda below.

1. **1st category: 77,420 total, 0-1145 per sector**
    - Planners, supervisors, leaders, murderers with excessive "zeal", those who committed rape or sexual torture
2. **2nd category: 430,552 total, 0-2807 per sector**
    - Those who killed, including those who killed unintentionally and those who attempted but failed
3. **3rd category: 309,645 total, 0-1919 per sector**
    - Those accused of property crimes

Those accused of the highest level of crimes are tried by the International Criminal Tribunal for Rwanda (ICTR) in Arusha. The rest are tried by Gacaca courts. The data released is the number of individuals accused of each of the three lower levels of crimes, in each of the 1513 sectors, who are still alive and living in Rwanda.[5] The detailed and disaggregated nature of the data provide obvious benefits, and the information about the number of perpetrators, rather than another measure of conflict intensity, is particularly appropriate for this project.

Still the data has limits. The most important is that it is not a measure of those who participated but a measure of those who are being prosecuted. There are a few ways in which these numbers could differ. For an individual to be included on these lists, somebody else in the area must have publicly claimed that he or she participated in the genocide. First, it is possible, and there are allegations of specific instances, that these accusations can be used politically or to settle old scores, and false accusations can be made. Conversely, it is very possible that many who participated were never accused, perhaps because all of the witnesses and friends and relatives of those they killed are not around to tell the story or are afraid to speak up. This bias is likely to be stronger than the first and thus we would expect

[4] A previous categorization scheme contained four levels of crimes, splitting the second category into two separate levels of crimes.

[5] Although not reported, these records are highly correlated with other measures of intensity of conflict based on either fewer numbers of communes or more aggregated administrative areas. These include a 1996 report from the Ministry of Higher Education, Scientific Research, and Culture (Kapiteni (1996)); the PRIO/Uppsala data on violent conflicts (Gleditsch et al. (2002)); and a database of timing and lethalty of conflict from Davenport and Stam (2009).

that in areas where there are fewer survivors, the number of accusations may be differentially lower than the number of participants. Still, the Gacaca courts have been very thorough in investigating, and reports of those afraid to speak are rare, so this data is likely to be a good proxy for the number of participants in each area. Specific concerns will be addressed in the analysis.

I aggregate these sector-level participation numbers to the level of a commune and combine them with the population by commune from the 1991 census. The population numbers are likely to be a very good approximation of the 1994 population. This is because 1991 is only three years earlier so population change from births is likely consistent across regions and also largely irrelevant because the new additions would be too young to participate. Mobility, the other source of population change, within the country was limited, requiring official approval to change residences (Prunier (1995)), and heavy restrictions on buying and selling land (André and Platteau (1998)). One possible alternative measure would be to use the population of adult men as the denominator, instead of the entire population, since perpetrators were nearly exclusively from this group. This, however, would require choosing an appropriate cut-off for ages and ignoring the small number of women who did participate, particularly at lower levels of crimes. Using the entire population reduces any threat of bias from the researcher choosing the cut-off. I will repeat the analysis using men age 18-50 as the denominator to be sure that the results do not qualitatively change. Figure 3.1 presents the observed variation in participation rates across communes.

3.4.2 Baseline Local Economic Conditions

These are constructed using the 1991 Census available through IPUMS International.

Ethnicity: The 1991 Census includes a variable for ethnicity. Before the genocide, all citizens were required to carry an identity card in which his or her ethnicity was clearly marked. In general, children of mixed marriages took the ethnicity of their father, but there are a few exceptions to this rule that show up in the Census (less than five percent of those whose parents are of different ethnicities). I limit the sample to only those who are either Hutu or Tutsi. The remainder of the observations in the Census includes the approximately one percent of the population that is Twa, naturalized Rwandans from other origins, and foreigners living in the country. Ignoring the others does not significantly change population numbers.

Human Capital: The human capital variable is presented as a score based on self-reports of the highest level of education completed. The codes for values of this variable were unavailable at higher levels. The lower numbers are coded to represent years of primary school and then years of secondary school, but values of the higher codes were hidden. This means that if these numbers were to be translated into years of schooling, then those with the most education would be dropped from the analysis. Therefore, the codes themselves are included directly as a human capital score. This is imperfect but does appear to reflect increases in human capital.

Employment: In an agricultural setting, virtually nobody reports being unemployed as those without formal sector jobs work on farms. For a measure of formal sector employment, I construct a binary variable that is one if the person reports that they are an employee or an employer, and zero if he or she reports being unemployed or self-employed. Using this,

I construct a measure of the percent of Hutu employed and the percent of Tutsi employed in the formal sector. This is similar to other measures of "off-farm income" used in the literature.

Assets: The census includes categorical variables for characterization of a set of physical assets (number of rooms per person; wall, floor, and roof materials; water, light and fuel sources; and type of toilet). Unfortunately, these measures can only generate a very rough measure of wealth or income in a way that reflects economic opportunity. For this reason, the analysis relies more heavily on employment.

Density: The census separately reports the area of each commune, which I use with the raw data to construct the population density.

Proportion Hutu: I construct a variable to represent the proportion of the population in a commune that is Hutu. In a country with two ethnic groups and enough tension between them to facilitate such a bloody history, the two ethnic groups are surprisingly dispersed which generates a proportion Hutu that ranges from 0.65 to 0.99. One explanation for the dispersion of Tutsi is that animal-keeping, the historical occupation of many Tutsi, is an offshoot of agriculturalist societies and depends on proximity to agriculturalists for nutrients. While animals provide milk, and occasionally meat, most herding societies use these products to trade for a dominant portion of their calories. Regardless of the origin, the current ratio of the ethnic groups can change the incentives of potential participants. As the proportion Tutsi increases, the potential reward of expropriation increases. However, if it gets high enough, the likelihood of success declines. Thus we expect, and find, that the highest level of participation comes when there are sufficient Tutsi for participation to be profitable. It should be noted that this relationship may be somewhat mechanical as one would expect there to be low participation when there is simply nobody to attack. Figure 3.2 plots participation rates as a function of the proportion of the population that is Hutu.

Young men: Participants in violent conflict are often young men. Young men are also likely to be relatively more educated with relatively fewer assets when compared with their elders. Thus if there is variation in the age distribution in communes, this could be an important omitted variable. We may also expect that different populations due to temporary migration may have different impacts on participation rates. I construct four variables to control for their presence: 1) The proportion of the population that is between 18 and 30, 2) The proportion of the population that is male, 3) the proportion of the population currently in the commune which was born elsewhere, and 4) the proportion of the population that was born in the commune that currently lives somewhere else.

There may be reason to think that the young male population was not as important in this conflict as in others. Based on first-hand accounts, all men were asked to participate, and where there were high levels of participation, men of all ages joined the ranks. Figure 3.3 presents the age distribution in Kibuye Prefecture of those in the Census and the victims. While we would not expect the age distribution of those who participated to match that of the victims, the similarities in the distribution may still support a claim that this war was not one simply between young men. Also, there are many accounts in which women participated. They were not generally taking part in killing squads, but they were often responsible for looting and were sometimes credited with encouraging their male relatives (Hatzfeld (2005), Hatzfeld (2013)).

Radio: Recent research has convincingly shown that the presence of radios contributed

enormously to the number of participants in the genocide as propaganda on the radio coordinated individuals (Yanagizawa-Drott (2010)). As radio-ownership may be associated with education and resources, it is included as a control variable. The census asks individuals whether their household owns a radio and the commune-level variable is the percent of Hutu and the percent of Tutsi who own a radio. These are included separately since radio ownership among Hutu is what has been previously associated with violence.

Table 5 presents a selection of summary statistics, split by ethnicity.

3.4.3 Combining Data

The administrative boundaries in Rwanda have shifted multiple times since the 1991 census, which means that matching the two data sources is not perfectly straight-forward. Gacaca records are reported by current prefecture and district, and by the pre-2005 sector. Census records are reported by province and commune, and limited information about population is presented with the sector. The analysis requires the use of information only available in the census at the commune level, and thus the sectors in the Gacaca records need to be matched with their former communes. The sectors correspond reasonably well, but different spellings and multiple sectors in different parts of the country with the same name complicate the process. The method used took two stages. In the first stage, sectors were matched automatically based on names and, in the absence of evidence to the contrary, a sector in one data set with a single correspondingly named sector in the second dataset was considered matched. In the second stage, the set of sectors that shared names in either dataset or that did not match a name in the other dataset were matched by hand. This involved using 1) the locations of the previous eleven provinces and the current four prefectures, 2) records from ICTR cases in which home sectors and communes of the accused were listed, and 3) any other mentions of the sector and commune and province available on the Internet. In the end, nearly 95 percent of sectors were matched. In cases in which sectors were not matched, I assume the participation rate in the unmatched sectors is equivalent to that in the matched sectors within the same commune.

3.5 Empirical Strategy

The general form of the estimating equation will be as follows:

$$ParticipationRate_j = a + b_1 * Employment(Tutsi)_j + b_2 * Employment(Hutu)_j \\ + b_3 * MeanEducation(Tutsi)_j + b_4 * MeanEducation(Hutu)_j + b_{ij} * X_{ij} + e_j$$

where $ParticipationRate_j$ is the participation rate measured by the number prosecuted divided by the Hutu population in a commune. The variable $Employment(Tutsi)_j$ represents the percent of Tutsi who are formally employed, while $MeanEducation(Tutsi)_j$ represents the mean education score for individuals in commune j. Finally, X_{ij} is a vector of controls. Thus the empirical predictions are as follows:

- If the *opportunity cost* theories are true, then we would expect b_2 and b_4 to be negative.

- If the *relative deprivation and expropriation* theories are true, then we would expect b_1 and b_3 to be positive.
- if the theories of violence as *political participation* are true, then b_4 may be positive.

Due to data constraints, I estimate all equations at the commune level, with 145 communes. A few variables are available at the sector level, including participation. It would be possible to use the sector as the unit analysis and cluster the standard errors at the level of the commune, but as the key independent variables are not available at the level of the sector, this would not change the statistical power. All equations are estimated controlling for density, the proportion Hutu, and the square of the proportion Hutu, the population of the commune, and province fixed effects for each of the 11 provinces.[6] In one specification, I include a commune-level mean asset-score for each ethnicity to broaden the story about employment. I also present estimates that control for the proportion of the population that is young and the proportion of the population that is male, as well as for the proportion that immigrated and that emigrated, the presence of radios among each ethnic group, and with all controls in the same specification. In all specifications, I aggregate the individual data by using the mean of the individual variables within a commune. Results are reported using all levels of crimes as outcomes, as well as using only the first two and only the third levels as the outcome. [7]

3.6 Results and Discussion

3.6.1 Estimates

Table 3.2 presents estimates using the overall participation rate as a percentage (0 to 100) as the dependent variable. All estimates include controls for the population of the commune, the proportion of the population that is Hutu, and the square of this, as well as the population density, and a full set of province fixed effects. Column 1 contains the basic specification with only these controls and education and employment for each ethnic group. The coefficient on Hutu education is positive and significant. The standard deviation of mean human capital scores at the commune level is 1.9, so a one standard deviation increase in the communes level of education is associated with an increase in the participation rate of 2.5 percentage points. The mean participation rate is 14 percent and the average Hutu population in a commune is approximately 45,000, so this represents an increase in participation of 18 percent or approximately 1125 participants per commune. The coefficient on Hutu employment is -44.09, which is significant at the 1 percent level. Here a one standard deviation change in the commune's employment rate would generate a reduction in the participation rate of 4 percentage points, a change of 29 percent and approximately 1800 participants per commune. The coefficients for human capital and employment for Tutsi are both small and insignificant.

[6]There are now only 5 provinces, but there were 11 at the time of the genocide in 1994 and in 1991 when the census was collected.

[7]All estimates are also repeated using the medians in place of the means or weighting each observation by the population of the commune, which do not qualitatively change the results.

Column 2 controls for the population that is young and the proportion of the population that is male. One concern is that young people have higher education, lower assets, and a higher propensity to participate in violence and thus their presence could independently generate similar correlations. Neither of these is statistically significant, but the estimates are sufficiently imprecise that it is not possible to rule out some meaningful relationship with the participation rate. Interestingly, their inclusion does not substantially change the strong and statistically significant coefficients on Hutu human capital and employment. The third column controls for the proportion of the population that was born elsewhere and the proportion born there which had emigrated from the commune. Neither of these is statistically significant and their inclusion also does not change the strength of the previously identified coefficients. The fourth column includes controls for radio ownership among each ethnic group and again, the main results do not change. The final column includes all controls, and again, the results remain stable.

This shows consistent estimates of the relationship between the participate-rate and Hutu education and employment. The coefficients on Tutsi human capital and employment are consistently of small magnitude and not statistically significant.

Tables 3.3 and 3.4 present the same estimates but split between types of crime. The participation rates used as the independent variable in Table 3.3 include only crimes in categories 1 and 2, which are necessarily violent, and Table 3.4 includes crimes that may have only dealt with property. Because the participation rate is the sum of these two types of crimes, the coefficients in Tables 3.3 and 3.4 sum to give the coefficients in Table 3.2, thus they are necessarily of smaller magnitude. The first two categories make up 62 percent of the total number of participants. The results in these two tables appear quite similar. One interpretation of the similarity is that the same motivations generate the two types of crimes. Alternatively, the two types of crimes were very correlated within the genocide and more violence may have made more property crime possible.

3.6.2 Discussion

Opportunity Costs

Opportunity cost theories predict that communes in which potential aggressors had the greatest level of available resources outside of conflict would be the least likely to have high participation rates. The consistently large negative coefficient on Hutu employment fits very well with this theory. The level of employment in an area reflects more than whether potential participants have jobs. It also reflects the degree to which labor markets function and the degree to which an individual can anticipate that they would be able to become employed. Such hope of a job would presumably be hurt by participation, while in areas where this hope does not exist, there is less to lose. Thus, while these findings cannot disentangle between the impact of actual and anticipated employment, they suggest an important role for opportunity costs in changing participation rates.

Repeating the analysis with a score for assets in place of employment (available upon request) generates qualitatively similar results, but with less precision due to measurement error in the asset measure.

Expropriation and Relative Deprivation

Expropriation and relative deprivation theories predict that the more resources that are held by potential victims of an attack, the more likely that attack is to occur. In this case, these would predict that communes with high levels of employment, and education among Tutsi would see the highest levels of violence. In Tables 3.2, 3.3, and 3.4, these relationships are all consistently insignificantly different from zero. It is important to note that this lack of a significant result does not conflict with other findings that - within a single area - the most likely to be killed are those that are well-off (Bundervoet (2009), André and Platteau (1998), De Walque and Verwimp (2010)), as the results are at the commune level.[8]

This paper then does not provide evidence in favor of the theories of looting and relative deprivation generating violence, although it also cannot conclusively say that these factors have no effect.

Political Participation and Unmet Expectations

The unique prediction of the theories of violence as political participation was that education among potential aggressors can increase the likelihood of violence. With a robust significant positive estimate of the relationship between Hutu education and participation, this paper provides evidence that is consistent with these theories.

Previous studies that have found a link between education and violence are often threatened by selection into violence. Terrorist organizations may choose the most educated members or those who have the most education within an area may find themselves leading action, even if their support for it is not greater than those they lead. In the Rwandan case, these threats are not such worries. First, everybody was encouraged to participate and thus selection is much less of a concern than in any context in which fighters are screened by the organization for which they are fighting. Second, this study relies on the aggregation of individual observations to the level of the commune, which can complicate the findings of some results, but in this case, it may strengthen them. The findings do not imply that those who are the most educated within their areas are the most likely to fight, but that in those areas where many are educated, some individuals are more likely to fight.

3.6.3 Further Analysis

Many of the theories that predict a positive association between education and violence include a link with unemployment or other factors that lead to dissatisfaction. If this is true, then one should expect that education and unemployment do not act separately to

[8]Classical measurement error may pose another threat to conclusions about the coefficients on Tutsi resources. In some communes, the number of Tutsi households is low enough that the mean levels are estimated by aggregating over a small number of observations. This means that the imprecise estimate of the local level of employment and education, especially in areas with small Tutsi populations, could generate classical measurement error in the independent variable and resulting attenuation bias in the coefficients. To check whether this is the case, I used an instrumental variables specification to reduce measurement error. This was done by splitting the sample of Tutsi households in half in each commune and using the estimated mean in one half as an instrument for the mean in the other half. These estimates, which are not presented, provide similarly insignificant coefficients on Tutsi employment and human capital.

influence violence but instead interact. To test whether this is true, I interact the two using two different methods.

In the first specification, the two variables are interacted at the local level. Unemployment is used instead of employment for ease of interpretation of the coefficients. This is defined as one minus the employment rate previously used. Then both this variable and the variable for Hutu education are normalized to have a mean of zero and a standard deviation of one. This is done so that the un-interacted terms can be interpreted as the effect at the mean of a change of one standard deviation. Then the usual specifications are repeated including the normalized unemployment and education variables and the interaction between the two. If education and unemployment combine to generate greater violence, then one expects a positive coefficient on the interaction term. This is presented in Table 3.5.

The coefficients on the un-interacted terms are still strong and significant in the same directions as before. The interaction term has a positive coefficient, which is only statistically significant in two of the five models. Still the magnitude of the coefficients is large. If we choose the coefficient from column 5, the coefficient is 0.665, which means that a one standard deviation increase in education increases the impact of unemployment on participation by 0.665 percentage points or 292 individuals.

The second specification looks at differences in education levels of individuals by employment status. Although the data must be aggregated to the level of the commune, the rich census data allows for the inclusion of many more statistics, beyond simple means. In this specification, the mean education among Hutu is included separated for those who are formally employed and those who are not formally employed. This is presented in Table 3.6. Based on the coefficients, the predictive power is largely driven by the mean education of those who are unemployed, rather than those with formal-sector employment. An F-test of the difference between these two coefficients is provided in each column. The differences are nearly significant at the ten percent level, but not quite.

3.7 Conclusion

This paper provides evidence that is consistent with models of violence as political participation and with opportunity costs reducing participation, but is not consistent with theories of relative deprivation and looting as key determinants of violence. A few caveats must be presented along with these findings.

First, this is only one case. Cramer (2003) pointed out that in the case of inequality in violence, one important explanation for variation in findings across conflicts or inconclusive cross-national results is that the causes of violence are extremely complex and that patterns that do exist may be stronger or weaker in different contexts. Still, a study of the variation in violence at the sub-national level contributes to a growing story about causes of conflict broadly. This case is also particularly interesting and broadly relevant. In only 100 days, nearly one million individuals were killed, mostly with tools used in farming. A similar number of people participated in this violence. The incredible magnitude of this cannot be ignored and it has understandably piqued the interest of a broad range of scholars attempting to understand or come to terms with it. Its applicability may also come from the scope and the nature of the participants. Very few or no other conflicts involved as many (partially)

voluntary participants and thus this provides an obvious case in which to investigate causes of individual participation.

Second, the aggregation to the level of the commune introduces the possibility of ecological fallacy, because individual actions cannot be linked to individual circumstances. This may call for a broader interpretation. For example, the results for employment may reflect an impact of living in an area with a functioning labor market in which there is the possibility of a job, even if the individual is not employed at the moment. This hope of employment based on substantial employment among co-ethnics could be what drives the observed difference in behavior.

This paper contributes to a greater understanding of the causes of violence by examining local economic conditions associated with variation in participation rates during the Rwandan genocide. It exploits recently released data on crimes during the genocide, in combination with individual-level economic information from before the violence. The findings - that increases in education and decreases in employment among potential attackers are associated with greater levels of violence and the resources of potential victims are not associated with violence levels - are consistent with opportunity cost theories of violence and theories of violence as political participation, but not with theories of relative deprivation and looting. Although still limited, to my knowledge, this is the best available data to quantitatively address causes of participation in the Rwandan genocide, and one of the best sources of sub-national data on civil conflict. This paper provides a new addition to the still limited set of quantitative, sub-national studies of violence.

The findings about a link between violence and the combination of relatively high levels of education with high levels of unemployment fit within a much larger discussion of the origins of participation in rebellions and political violence, in particular the discussion surrounding the Arab Spring. This leads to two different conclusions. First, it provides an argument in favor of recognizing that participation in the genocide in Rwanda ought to be conceived of as a form of political violence and not simply [arbitrary/not thought out] violence. Second, his provides an additional piece of evidence to confirm that this link exists, showing that education alone is not enough to create a peaceful society, and it encourages much more research and investment into creating employment for the enormous number of newly educated young people in the developing world.

3.8 Figures

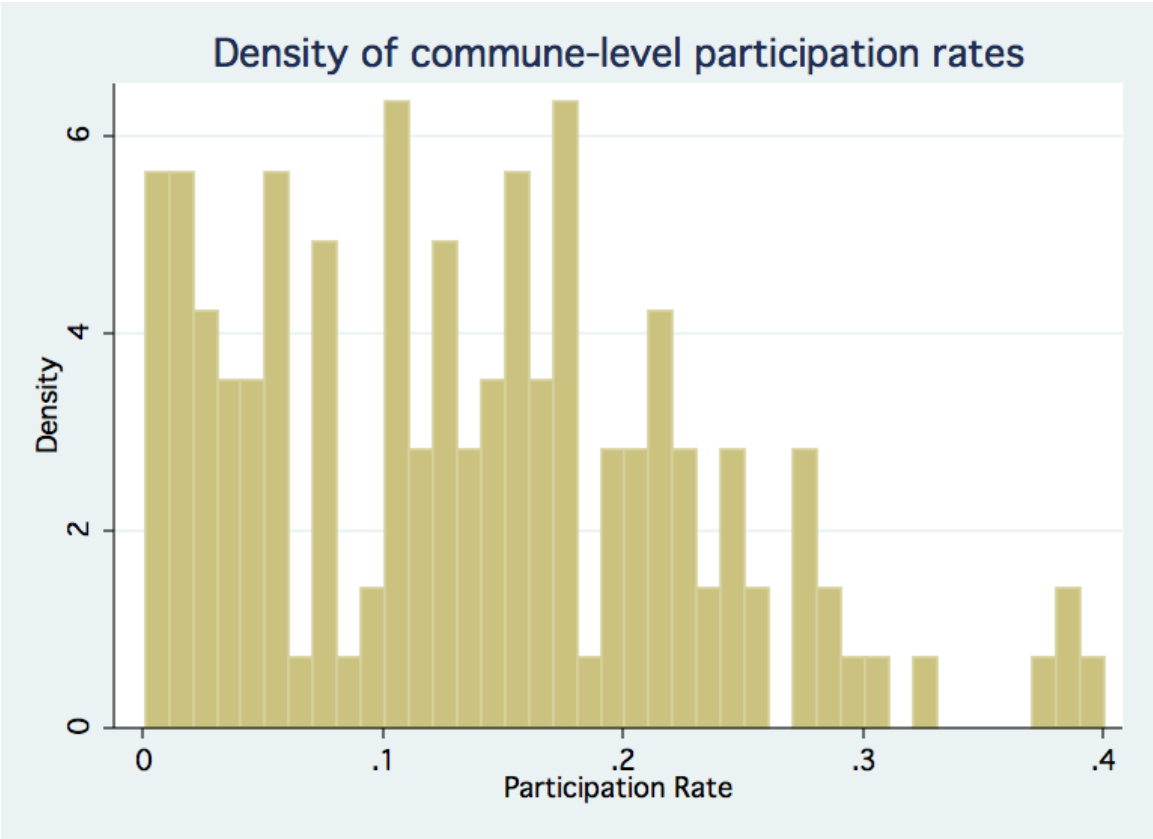


Figure 3.1: This is the distribution of participation rates at the commune level. The participation rate is measured as the number of people accused in the Gacaca records, divided by the Hutu population in 1991 from the census.

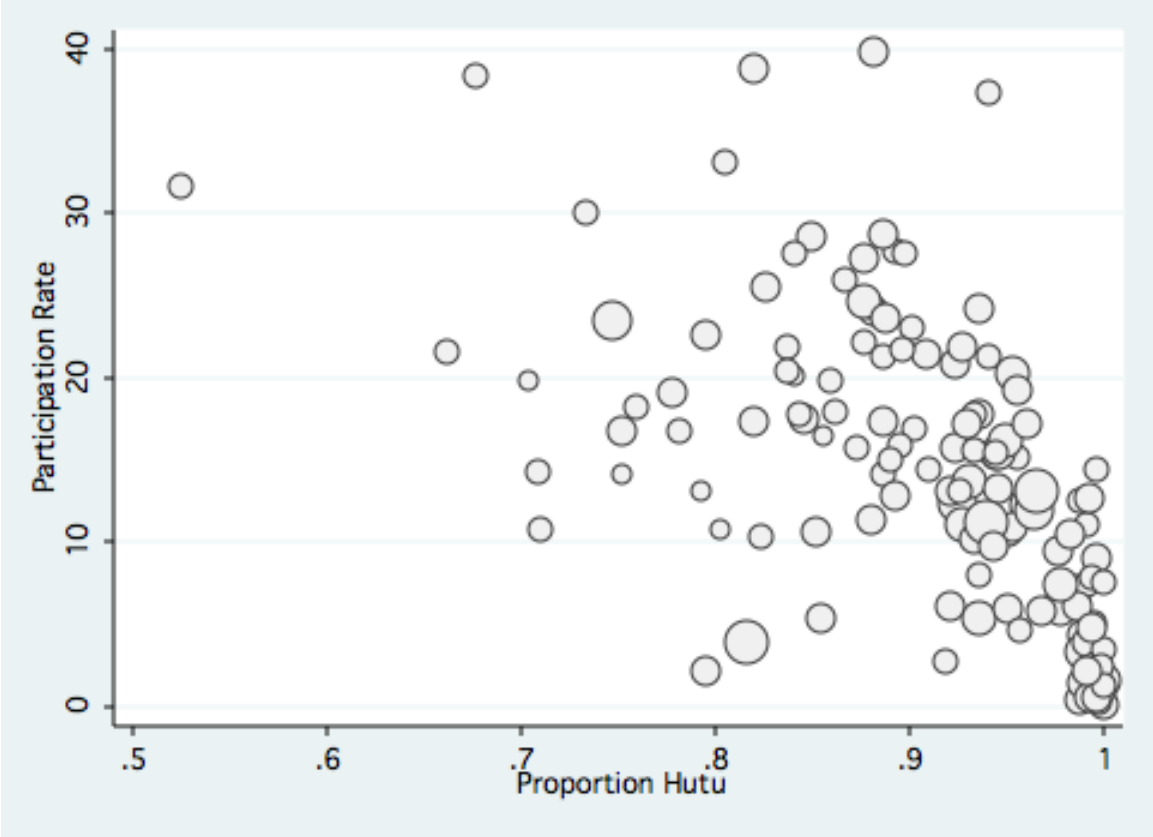


Figure 3.2: Commune-level participation rates as a function of the proportion Hutu. Each dot represents one commune, and the size is in proportion to the population of that commune.

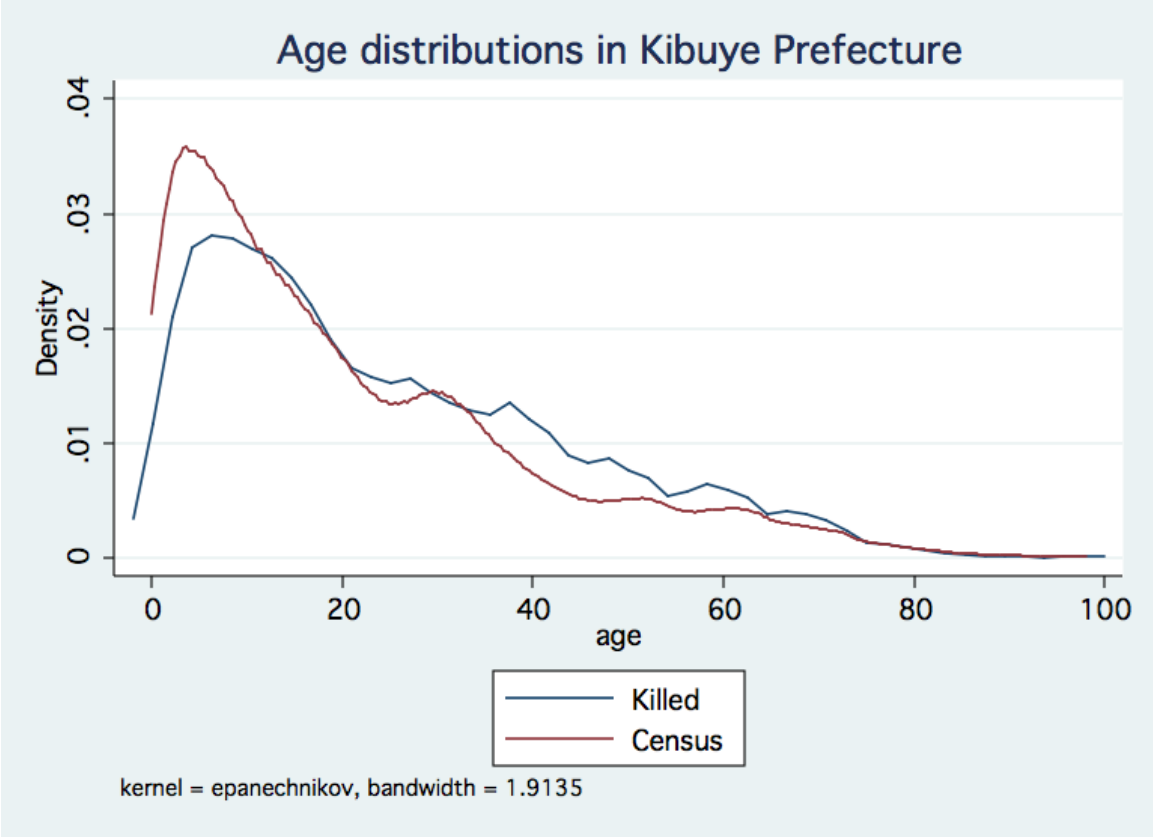


Figure 3.3: This plots the distribution of ages reported to have been killed in the Verwimp (2004) data from Kibuye on top of the distribution of all ages reported in the 1991 census.

3.9 Tables

Table 3.1: Summary Statistics

Variable	Hutu	Tutsi
Population share	0.871	0.081
	(0.335)	(0.273)
Human Capital Score	9.313	12.818
	(8.910)	(9.789)
Literate	0.517	0.673
	(0.500)	(0.469)
Employed	0.074	0.113
	(0.262)	(0.317)
Asset Score	-0.024	0.214
	(0.920)	(1.279)
Owns a radio	0.286	0.349
	(0.452)	(0.477)
Has electricity	0.021	0.046
	(0.143)	(0.210)
Owns land	0.928	0.905
	(0.258)	(0.294)
Observations	647010	60096

Means from 1991 Census data. Standard deviations in parentheses.

Table 3.2: Basic specifications

VARIABLES	(1) part_rate	(2) part_rate	(3) part_rate	(4) part_rate	(5) part_rate
Mean Education (Tutsi)	-0.0597	-0.0850	-0.0644	0.0752	0.0497
	(0.0941)	(0.0976)	(0.107)	(0.118)	(0.128)
Mean Education (Hutu)	1.331**	1.202**	1.488**	1.253**	1.302*
	(0.523)	(0.536)	(0.603)	(0.601)	(0.693)
Percent formally employed (Tutsi)	0.0491	0.399	0.226	-0.117	0.207
	(2.455)	(2.651)	(2.916)	(2.216)	(2.703)
Percent formally employed (Hutu)	-44.09***	-49.84***	-46.20***	-43.65***	-48.24***
	(11.50)	(12.83)	(12.69)	(11.26)	(12.59)
Percent 18-30		40.99			37.29
		(37.23)			(41.57)
Percent male		14.77			4.498
		(64.26)			(70.85)
Percent born away			-1.716		1.351
			(7.038)		(8.299)
Percent left			13.07		13.03
			(12.61)		(12.96)
Radio ownership (Tutsi)				-3.738*	-3.951*
				(2.139)	(2.257)
Radio ownership (Hutu)				1.233	1.872
				(10.83)	(11.49)
Observations	137	137	137	135	135
R-squared	0.689	0.692	0.692	0.684	0.689

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
All specifications include controls for density, population, proportion Hutu, and provincial fixed effects

Table 3.3: Only categories 1 and 2

VARIABLES	(1) part_rate12	(2) part_rate12	(3) part_rate12	(4) part_rate12	(5) part_rate12
Mean Education (Tutsi)	-0.0417	-0.0658	-0.0427	0.0302	0.0145
	(0.0544)	(0.0593)	(0.0652)	(0.0674)	(0.0773)
Mean Education (Hutu)	0.712**	0.615*	0.809**	0.670*	0.692
	(0.329)	(0.337)	(0.383)	(0.382)	(0.439)
Percent formally employed (Tutsi)	-0.0259	0.390	0.0967	-0.157	0.154
	(1.347)	(1.566)	(1.728)	(1.302)	(1.708)
Percent formally employed (Hutu)	-22.11***	-27.79***	-25.38***	-21.60***	-27.37***
	(7.178)	(8.648)	(8.159)	(6.718)	(8.226)
Percent 18-30		28.69			21.98
		(24.53)			(28.06)
Percent male		27.37			14.90
		(44.17)			(47.54)
Percent born away			-3.984		-1.936
			(5.361)		(6.123)
Percent left			10.06		9.506
			(8.310)		(8.637)
Radio ownership (Tutsi)				-2.111	-2.274
				(1.369)	(1.450)
Radio ownership (Hutu)				0.478	0.918
				(7.038)	(7.264)
Observations	137	137	137	135	135
R-squared	0.687	0.692	0.693	0.681	0.690

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

All specifications include controls for density, population, proportion Hutu, and provincial fixed effects

Table 3.4: Only category 3

VARIABLES	(1) part_rate3	(2) part_rate3	(3) part_rate3	(4) part_rate3	(5) part_rate3
Mean Education (Tutsi)	-0.0180	-0.0192	-0.0217	0.0450	0.0352
	(0.0477)	(0.0462)	(0.0499)	(0.0641)	(0.0640)
Mean Education (Hutu)	0.619***	0.587**	0.679**	0.583**	0.610*
	(0.235)	(0.243)	(0.266)	(0.267)	(0.310)
Percent formally employed (Tutsi)	0.0750	0.00834	0.129	0.0406	0.0526
	(1.244)	(1.224)	(1.308)	(1.083)	(1.148)
Percent formally employed (Hutu)	-21.98***	-22.05***	-20.82***	-22.05***	-20.88***
	(5.674)	(5.658)	(5.965)	(6.281)	(6.452)
Percent 18-30		12.31			15.31
		(17.02)			(18.43)
Percent male		-12.60			-10.40
		(27.55)			(30.41)
Percent born away			2.268		3.286
			(2.519)		(3.011)
Percent left			3.003		3.526
			(5.688)		(5.786)
Radio ownership (Tutsi)				-1.627	-1.677
				(1.034)	(1.056)
Radio ownership (Hutu)				0.755	0.954
				(5.550)	(5.914)
Observations	137	137	137	135	135
R-squared	0.627	0.629	0.629	0.623	0.627

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
All specifications include controls for density, population, proportion Hutu, and provincial fixed effects

Table 3.5: Interaction of education and unemployment within commune

VARIABLES	(1) part_rate	(2) part_rate	(3) part_rate	(4) part_rate	(5) part_rate
Hutu Mean Ed (normalized)	2.665***	2.171**	2.711**	2.326**	2.285*
	(1.004)	(1.004)	(1.128)	(1.140)	(1.289)
Hutu Unemployment (normalized)	3.351***	3.548***	3.365***	3.075***	3.500***
	(1.080)	(1.185)	(1.186)	(1.055)	(1.151)
Mean Ed*Unemployment (Hutu)	0.279	0.703*	0.525	0.532	0.665*
	(0.226)	(0.362)	(0.372)	(0.352)	(0.385)
Mean Education (Tutsi)	-0.0540	-0.0802	-0.0525	0.0882	0.0581
	(0.0932)	(0.0972)	(0.106)	(0.112)	(0.121)
Percent formally employed (Tutsi)	-0.0690	-0.0867	-0.268	-0.566	-0.301
	(2.508)	(2.753)	(2.993)	(2.218)	(2.737)
Percent 18-30		51.22			46.22
		(38.22)			(41.91)
Percent male		23.68			14.16
		(64.95)			(71.09)
Percent born away			-3.240		0.350
			(7.434)		(8.506)
Percent left			11.72		10.79
			(12.60)		(13.04)
Radio ownership (Tutsi)				-3.800*	-3.980*
				(2.057)	(2.165)
Radio ownership (Hutu)				1.366	2.075
				(10.82)	(11.45)
Observations	137	137	137	135	135
R-squared	0.689	0.697	0.694	0.686	0.693

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

All specifications include controls for density, population, proportion Hutu, and provincial fixed effects

Table 3.6: Interaction of education and unemployment within individual

VARIABLES	(1) part_rate	(2) part_rate	(3) part_rate	(4) part_rate	(5) part_rate
Mean Education (Tutsi)	-0.0597	-0.0820	-0.0639	0.0817	0.0601
	(0.0967)	(0.0998)	(0.109)	(0.124)	(0.131)
Mean Ed of Unemployed Hutu	1.220**	1.088*	1.309**	1.165*	1.149
	(0.598)	(0.610)	(0.656)	(0.654)	(0.725)
Mean Ed of Employed Hutu	0.0980	0.0856	0.122	0.0787	0.0856
	(0.188)	(0.192)	(0.192)	(0.197)	(0.205)
Percent formally employed (Tutsi)	-0.0987	0.241	0.0403	-0.240	0.0629
	(2.494)	(2.706)	(2.943)	(2.200)	(2.671)
Percent formally employed (Hutu)	-28.92***	-35.77***	-28.95***	-30.59***	-34.66**
	(9.750)	(12.19)	(10.46)	(11.35)	(13.39)
Percent 18-30		37.67			33.58
		(37.22)			(41.81)
Percent male		16.29			8.553
		(64.57)			(71.11)
Percent born away			-1.715		1.357
			(7.155)		(8.494)
Percent left			11.98		12.09
			(12.49)		(12.90)
Radio ownership (Tutsi)				-3.834*	-4.081*
				(2.175)	(2.295)
Radio ownership (Hutu)				2.313	3.483
				(10.71)	(11.36)
Observations	137	137	137	135	135
R^2	0.689	0.692	0.691	0.684	0.689
F-stat	2.663	2.040	2.598	2.188	1.783
p-value	0.105	0.156	0.110	0.142	0.185

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
All specifications include controls for density, population, proportion Hutu, and provincial FEs.

Bibliography

Abadie, Alberto. 2006. "Poverty, Political Freedom, and the Roots of Terrorism." *American Economic Review*, 96(2): 50–56.

Alesina, Alberto, and Roberto Perotti. 1996. "Income distribution, political instability, and investment." *European Economic Review*, 40(6): 1203–1228.

André, Catherine, and Jean-Philippe Platteau. 1998. "Land relations under unbearable stress: Rwanda caught in the Malthusian trap." *Journal of Economic Behavior & Organization*, 34(1): 1–47.

Auvert, Bertran, Dirk Taljaard, Emmanuel Lagarde, Joelle Sobngwi-Tambekou, Rémi Sitta, and Adrian Puren. 2005. "Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial." *PLoS medicine*, 2(11): e298.

Bailey, Robert C, Stephen Moses, Corette B Parker, Kawango Agot, Ian Maclean, John N Krieger, Carolyn FM Williams, Richard T Campbell, and Jeckoniah O Ndinya-Achola. 2007. "Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial." *The Lancet*, 369(9562): 643–656.

Bandiera, Oriana, Andrea Prat, and Tommaso Valletti. 2009. "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." *The American Economic Review*, 99(4): 1278–1308.

Banerjee, Abhijit, and Esther Duflo. 2006. "Addressing absence." *The journal of economic perspectives: a journal of the American Economic Association*, 20(1): 117.

Banerjee, Abhijit, Donald P Green, Jeffery McManus, and Rohini Pande. 2012. "Are Poor Voters Indifferent to Whether Elected Leaders are Criminal or Corrupt? A Vignette Experiment in Rural India."

Banerjee, Abhijit V, Esther Duflo, and Rachel Glennerster. 2008. "Putting a Band-Aid on a corpse: Incentives for nurses in the Indian public health care system." *Journal of the European Economic Association*, 6(2-3): 487–500.

Baranov, Victoria, and Hans-Peter Kohler. 2012. "The impact of AIDS treatment on savings and human capital investment in Malawi." Vol. 3.

Baranov, Victoria, Daniel Bennett, and Hans-Peter Kohler. 2012. "The Indirect Impact of Antiretroviral Therapy."

Bardhan, Pranab. 1997. "Corruption and development: a review of issues." *Journal of Economic Literature*, 35(3): 1320–1346.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2008. "Adult mortality and consumption growth in the age of HIV/AIDS." *Economic Development and Cultural Change*, 56(2): 299–326.

Bell, Clive, Shantayanan Devarajan, and Hans Gersbach. 2004. "Thinking about the long-run economic costs of AIDS." *The macroeconomics of HIV/AIDS*, 96: 128–129.

Berman, Eli, Michael Callen, Joseph H Felter, and Jacob N Shapiro. 2011. "Do working men rebel? Insurgency and unemployment in Afghanistan, Iraq, and the Philippines." *Journal of Conflict Resolution*, 55(4): 496–528.

Berrebi, Claude. 2007. "Evidence about the link between education, poverty and terrorism among Palestinians." *Peace Economics, Peace Science and Public Policy*, 13(1): 2.

Besley, Timothy, Rohini Pande, and Vijayendra Rao. Forthcoming. "Just Rewards? Local Politics and Public Resource Allocation in South India." *World Bank Economic Review*.

Besley, Timothy, Rohini Pande, Lupin Rahman, and Vijayendra Rao. 2004. "The politics of public good provision: Evidence from Indian local governments." *Journal of the European Economic Association*, 2(2-3): 416–426.

Blattman, Christopher, and Edward Miguel. 2010. "Civil war." *Journal of Economic Literature*, 3–57.

Blower, SM, Hayley B Gershengorn, and RM Grant. 2000. "A tale of two futures: HIV and antiretroviral therapy in San Francisco." *Science*, 287(5453): 650–654.

Bor, Jacob. 2012. "HIV Treatment and Labor Supply in Rural South Africa."

Boudreaux, Karol. 2009. "Land conflict and genocide in Rwanda." *The Electronic Journal of Sustainable Development*, 1(3): 86–95.

Bundervoet, Tom. 2009. "Livestock, land and political power: The 1993 killings in Burundi." *Journal of Peace Research*, 46(3): 357–376.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padro i Miquel. 2011. "Ethnic favoritism." *Unpublished ms*.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2011. "Robust inference with multiway clustering." *Journal of Business & Economic Statistics*, 29(2): 238–249.

Cavalcanti Ferreira, Pedro, and Samuel Pessoa. 2003. "The Long-Run Economic Impact of AIDS." *Available at SSRN 411782*.

Cederman, Lars-Erik, Andreas Wimmer, and Brian Min. 2010. "Why do ethnic groups rebel?: New data and analysis." *World Politics*, 62(1): 87–119.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F Halsey Rogers. 2006. "Missing in action: teacher and health worker absence in developing countries." *The Journal of Economic Perspectives*, 20(1): 91–116.

Chong, Alberto, Ana De La O, Dean S Karlan, and Léonard Wantchékon. 2012. "Looking Beyond the Incumbent: The Effects of Exposing Corruption on Electoral Outcomes." CEPR Discussion Papers.

Chua, Amy. 2004. *World on fire: How exporting free market democracy breeds ethnic hatred and global instability.* Anchor.

Cohen, Myron S, Ying Q Chen, Marybeth McCauley, Theresa Gamble, Mina C Hosseinipour, Nagalingeswaran Kumarasamy, James G Hakim, Johnstone Kumwenda, Beatriz Grinsztejn, Jose HS Pilotto, et al. 2011. "Prevention of HIV-1 infection with early antiretroviral therapy." *New England Journal of Medicine*, 365(6): 493–505.

Collier, Paul, and Anke Hoeffler. 2002. "On the incidence of civil war in Africa." *Journal of Conflict Resolution*, 46(1): 13–28.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and grievance in civil war." *Oxford economic papers*, 56(4): 563–595.

Corrigan, Paul, Gerhard Glomm, and Fabio Mendez. 2005. "AIDS crisis and growth." *Journal of Development Economics*, 77(1): 107–124.

Cramer, Christopher. 2003. "Does inequality cause conflict?" *Journal of International Development*, 15(4): 397–412.

Cuddington, John T, and John D Hancock. 1994. "Assessing the impact of AIDS on the growth path of the Malawian economy." *Journal of Development Economics*, 43(2): 363–368.

Dal Bó, Ernesto, and Pedro Dal Bó. 2011. "Workers, warriors, and criminals: social conflict in general equilibrium." *Journal of the European Economic Association*, 9(4): 646–677.

Davenport, Christian, and Allan Stam. 2009. "Rwandan political violence in space and time." *Unpublished manuscript.*

Davies, James Chowning. 1974. "The J-curve and power struggle theories of collective violence." *American Sociological Review*, 39(4): 607–610.

De Mesquita, Ethan Bueno. 2005. "The quality of terror." *American Journal of Political Science*, 49(3): 515–530.

Demombynes, Gabriel, and Berk Özler. 2005. "Crime and local inequality in South Africa." *Journal of Development Economics*, 76(2): 265–292.

Des Forges, Alison Liebhafsky, et al. 1999. *"Leave none to tell the story": genocide in Rwanda.* Vol. 3169, Human Rights Watch New York.

De Walque, Damien, and Philip Verwimp. 2010. "The demographic and socio-economic distribution of excess mortality during the 1994 genocide in Rwanda." *Journal of African Economies*, 19(2): 141–162.

de Walque, Damien, Harounan Kazianga, and Mead Over. 2012. "Antiretroviral Therapy Perceived Efficacy and Risky Sexual Behaviors: Evidence from Mozambique." *Economic Development and Cultural Change*, 61(1): 97–126.

Diamond, Jared. 2005. *Collapse: how societies choose to fail or succeed: revised edition.* Penguin.

Do, Quy-Toan, and Lakshmi Iyer. 2010. "Geography, poverty and conflict in Nepal." *Journal of Peace Research*, 47(6): 735–748.

Dube, Oeindrila, and Juan Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia*." *The Review of Economic Studies*.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Education, HIV and early fertility: Experimental evidence from Kenya." *UCLA manuscript*.

Dupas, Pascaline. 2011. "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics*, 3(1): 1–34.

Fajnzlber, Pablo, Daniel Lederman, and Norman Loayza. 2002. "Inequality and violent crime." *JL & Econ.*, 45: 1.

Fearon, James D, and David D Laitin. 2003. "Ethnicity, insurgency, and civil war." *American political science review*, 97(1): 75–90.

Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The effects of Brazil's publicly released audits on electoral outcomes." *The Quarterly Journal of Economics*, 123(2): 703–745.

Ferraz, Claudio, Frederico Finan, and Diana B Moreira. 2012. "Corrupting learning: Evidence from missing federal education funds in Brazil." *Journal of Public Economics*.

Filmer, Deon, Jeffrey S Hammer, and Lant H Pritchett. 2000. "Weak links in the chain: a diagnosis of health policy in poor countries." *The World Bank Research Observer*, 15(2): 199–224.

Fisman, Raymond, and Edward Miguel. 2007. "Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets." *Journal of Political Economy*, 115(6): 1020–1048.

Fortson, Jane G. 2011. "Mortality risk and human capital investment: The Impact of HIV/AIDS in Sub-Saharan Africa." *The Review of Economics and Statistics*, 93(1): 1–15.

Franck, Raphaël, and Ilia Rainer. 2012. "Does the Leaders Ethnicity Matter? Ethnic Favoritism, Education and Health in Sub-Saharan Africa."

Gertler, Paul, Manisha Shah, and Stefano M Bertozzi. 2005. "Risky business: the market for unprotected commercial sex." *Journal of Political Economy*, 113(3): 518–550.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. "Armed conflict 1946-2001: A new dataset." *Journal of peace research*, 39(5): 615–637.

Godlonton, Susan, Alister Munthali, and Rebecca Thornton. 2011. "Behavioral response to information? Circumcision, information, and HIV prevention." University of Michigan Working Paper.

Goldstein, Markus, Joshua Graff Zivin, James Habyarimana, Cristian Pop-Eleches, and Harsha Thirumurthy. 2013. "The Effect of Absenteeism and Clinic Protocol on Health Outcomes: The Case of Mother-to-Child Transmission of HIV in Kenya." *American Economic Journal: Applied Economics*, 5: 58–85.

Goldstein, Markus, Joshua Graff Zivin, James Habyarimana, Cristian Pop-Eleches, and Harsha Thirumurthyk. 2008. "Health worker absence, HIV testing and behavioral change: Evidence from western Kenya." *Unpublished Working Paper, Department of Economics, Columbia University*.

Gong, Erick. 2011. "HIV Testing and Risky Sexual Behavior." Middlebury College, Department of Economics.

Gourevitch, Philip. 1998. *We wish to inform you that tomorrow we will be killed with our families: Stories from Rwanda.* Picador USA.

Granich, Reuben M, Charles F Gilks, Christopher Dye, Kevin M De Cock, and Brian G Williams. 2009. "Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model." *The Lancet*, 373(9657): 48–57.

Gray, Ronald H, Godfrey Kigozi, David Serwadda, Frederick Makumbi, Stephen Watya, Fred Nalugoda, Noah Kiwanuka, Lawrence H Moulton, Mohammad A Chaudhary, Michael Z Chen, et al. 2007. "Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial." *The Lancet*, 369(9562): 657–666.

Grossman, Herschel I, and Minseong Kim. 1995. "Swords or plowshares? A theory of the security of claims to property." *Journal of Political Economy*, 1275–1288.

Gupta, Sanjeev, Hamid Reza Davoodi, and Erwin Tiongson. 2000. *Corruption and the provision of health care and education services.* International Monetary Fund.

Gurr, Ted R. 1970. *Why men rebel.* Princeton University Press.

Haavelmo, Trygve. 1954. *A study in the theory of economic evolution.* North-Holland Publishing Company.

Hatzfeld, Jean. 2005. *Machete season: The killers in Rwanda speak.* Farrar Straus & Giroux.

Hatzfeld, Jean. 2013. *Life Laid Bare: The Survivors of Rwanda Speak.* Other Press.

Hegre, Håvard, and Nicholas Sambanis. 2006. "Sensitivity analysis of empirical results on civil war onset." *Journal of Conflict Resolution*, 50(4): 508–535.

Hirshleifer, Jack. 1995. "Anarchy and its breakdown." *Journal of Political Economy*, 26–52.

Hodler, Roland, and Paul A Raschky. 2010. *Foreign Aid and Enlightened Leaders.* Studienzentrum Gerzensee.

Humphreys, Macartan, and Jeremy M Weinstein. 2008. "Who fights? The determinants of participation in civil war." *American Journal of Political Science*, 52(2): 436–455.

Hunter, Susan, and John Williamson. 2000. "Children on the Brink." *USAID, Washington DC*.

Huntington, Samuel P, and Francis Fukuyama. 2006. *Political order in changing societies.* Yale University Press.

Iyengar, Radha, Jonathan Monten, and Matthew Hanson. 2011. "Building Peace: The Impact of Aid on the Labor Market for Insurgents." *NBER Working Paper*, , (w17297).

Jablonski, Ryan. 2012. "How Aid Targets Votes: The Impact of Electoral Incentives on Foreign Aid Distribution."

Jamison, Julian, and Dean Karlan. 2011. "Measuring Preferences and Predicting Outcomes." 1–3.

Juhn, Chinhui, Sebnem Kalemli-Ozcan, and Belgi Turan. 2008. "HIV and fertility in Africa: first evidence from population-based surveys." *Journal of Population Economics*, 1–19.

Justino, Patricia. 2009. "Poverty and violent conflict: A micro-level perspective on the causes and duration of warfare." *Journal of Peace Research*, 46(3): 315–333.

Justino, Patricia, and Philip Verwimp. 2008. "Poverty dynamics, violent conflict and convergence in Rwanda."

Kalemli-Ozcan, Sebnem, and Belgi Turan. 2011. "HIV and fertility revisited." *Journal of Development Economics*, 96(1): 61–65.

Kaler, Amy. 2003. ""My girlfriends could fill a yanu-yanu bus: Rural Malawian mens claims about their own serostatus." *Demographic Research*, 19.

Kangwana, Beth B, Julius Njogu, Beatrice Wasunna, Sarah V Kedenge, Dorothy N Memusi, Catherine A Goodman, Dejan Zurovac, and Robert W Snow. 2009. "Malaria drug shortages in Kenya: a major failure to provide access to effective treatment." *The American journal of tropical medicine and hygiene*, 80(5): 737.

Kapiteni, Antoine. 1996. "La premire estimation du hombre des victims du gnocide du Rwanda de 1994 commune par commune en Fev 1996." *Report of the Ministry of Higher Education, Scientific Research, and Culture*.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2010. "The worldwide governance indicators: methodology and analytical issues." *World Bank policy research working paper*, , (5430).

Kelly, Morgan. 2000. "Inequality and crime." *Review of Economics and Statistics*, 82(4): 530–539.

Kerwin, Jason T. 2012. ""Rational Fatalism: Non-monotonic choices in response to risk."

Kopstein, Jeffrey S, and Jason Wittenberg. 2011. "Deadly communities: Local political milieus and the persecution of Jews in occupied Poland." *Comparative Political Studies*, 44(3): 259–283.

Kramon, Eric, and Daniel N Posner. 2012. "Ethnic Favoritism in Primary Education in Kenya." *Unpublished manuscript*.

Kremer, Michael. 1996. "Integrating behavioral choice into epidemiological models of AIDS." *The Quarterly Journal of Economics*, 111(2): 549–573.

Krueger, Alan B, and Jitka Maleckova. 2003. "Education, poverty and terrorism: Is there a causal connection?" *The Journal of Economic Perspectives*, 17(4): 119–144.

Kudamatsu, Masayuki. 2009. "Ethnic favoritism: Micro evidence from Guinea." *Available at SSRN 1440303*.

Lakdawalla, Darius, Neeraj Sood, and Dana Goldman. 2006. "HIV breakthroughs and risky sexual behavior." *The Quarterly Journal of Economics*, 121(3): 1063–1102.

Land, Kenneth C, Patricia L McCall, and Lawrence E Cohen. 1990. "Structural covariates of homicide rates: Are there any invariances across time and social space?" *American journal of sociology*, 922–963.

Law, Matthew G, Garrett Prestage, Andrew Grulich, Paul Van de Ven, and Susan Kippax. 2001. "Modelling the effect of combination antiretroviral treatments on HIV incidence." *Aids*, 15(10): 1287–1294.

Leff, Nathaniel H. 1964. "Economic development through bureaucratic corruption." *American Behavioral Scientist*, 8(3): 8–14.

Lewis, Maureen. 2006. "Governance and corruption in public health care systems." *Center for Global Development Working Paper*, , (78).

Lichbach, Mark Irving. 1989. "An evaluation of does economic inequality breed political conflict?studies." *World Politics*, 41(4): 431–470.

Lorentzen, Peter, John McMillan, and Romain Wacziarg. 2008. "Death and development." *Journal of Economic Growth*, 13(2): 81–124.

Lucas, Adrienne M, and Nicholas L Wilson. n.d.. "Adult Antiretroviral Therapy and Child Health: Evidence from Scale-up in Zambia."

Macours, Karen. 2011. "Increasing inequality and civil conflict in Nepal." *Oxford Economic Papers*, 63(1): 1–26.

Mamdani, Mahmood. 2002. *When victims become killers: Colonialism, nativism, and the genocide in Rwanda.* Princeton University Press.

Mason, T David, and Jerry A Murtagh. 1985. "Who riots? An empirical examination of the new urban black versus the social marginality hypotheses." *Political Behavior*, 7(4): 352–373.

Mauro, Paolo. 1995. "Corruption and growth." *The quarterly journal of economics*, 110(3): 681–712.

Mauro, Paolo. 1998. "Corruption and the composition of government expenditure." *Journal of Public economics*, 69(2): 263–279.

McLaren, Zoë M. 2012. "The Effect of Access to AIDS Treatment on Employment Outcomes in South Africa."

Mechoulan, Stéphane. 2007. "Risky Sexual Behavior, Testing and HIV Treatments."

Merton, Robert K. 1934. "Durkheim's division of labor in society." *American Journal of Sociology*, 319–328.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. "Economic shocks and civil conflict: An instrumental variables approach." *Journal of political Economy*, 112(4): 725–753.

Miller, Abraham H, Louis H Bolce, and Mark Halligan. 1977. "The J-curve theory and the black urban riots: an empirical test of progressive relative deprivation theory." *The American Political Science Review*, 71(3): 964–982.

Mills, Edward J, Jean B Nachega, Iain Buchan, James Orbinski, Amir Attaran, Sonal Singh, Beth Rachlis, Ping Wu, Curtis Cooper, Lehana Thabane, et al. 2006. "Adherence to antiretroviral therapy in sub-Saharan Africa and North America." *JAMA: the journal of the American Medical Association*, 296(6): 679–690.

Minnis, Alexandra M, Markus J Steiner, Maria F Gallo, Lee Warner, Marcia M Hobbs, Ariane van der Straten, Tsungai Chipato, Maurizio Macaluso, and Nancy S Padian. 2009. "Biomarker validation of reports of recent sexual activity: results of a randomized controlled study in Zimbabwe." *American journal of epidemiology*, 170(7): 918–924.

Mitra, Anirban, and Debraj Ray. 2010. "Implications of an economic theory of conflict." mimeo.

Murshed, S Mansoob, and Scott Gates. 2005. "Spatial–horizontal inequality and the Maoist insurgency in Nepal." *Review of Development Economics*, 9(1): 121–134.

Olken, Benjamin A. 2006. "Corruption and the costs of redistribution: Micro evidence from Indonesia." *Journal of Public Economics*, 90(4): 853–870.

Olken, Benjamin A. 2009. "Corruption perceptions vs. corruption reality." *Journal of Public Economics*, 93(7): 950–964.

Olken, Benjamin A, and Rohini Pande. 2012. "Corruption in Developing Countries." *Annual Review of Economics*, 4(1): 479–509.

Østby, Gudrun. 2008. "Polarization, horizontal inequalities and violent civil conflict." *Journal of Peace Research*, 45(2): 143–162.

Østby, Gudrun, Ragnhild Nordås, and Jan Ketil Rød. 2009. "Regional Inequalities and Civil Conflict in Sub-Saharan Africa1." *International Studies Quarterly*, 53(2): 301–324.

Oster, Emily. 2012. "HIV and sexual behavior change: Why not Africa?" *Journal of Health Economics*, 31(1): 35–49.

Over, Mead, Elliot Marseille, Kurapati Sudhakar, Julian Gold, Indrani Gupta, Abhaya Indrayan, Subhash Hira, Nico Nagelkerke, Arni SR Srinivasa Rao, and Peter Heywood. 2006. "Antiretroviral therapy and HIV prevention in India: modeling costs and consequences of policy options." *Sexually transmitted diseases*, 33(10): S145–S152.

Papageorge, Nick W. 2012. "HIV and Incentives for Monogamy among Gay Men."

Peltzman, Sam. 1975. "The effects of automobile safety regulation." *The Journal of Political Economy*, 677–725.

Percival, Val, and Thomas Homer-Dixon. 1998. "Environmental scarcity and violent conflict: The case of South Africa." *Journal of Peace Research*, 35(3): 279–298.

Posner, Daniel, and Eric Kramon. 2011. "Who benefits from distributive politics? How the outcome one studies affects the answer one gets."

Pritchett, Lant. 1996. "Mind your p's and q's: the cost of public investment is not the value of public capital." *World Bank policy research working paper*, , (1660).

Prunier, Gérard. 1995. *The Rwanda crisis: History of a genocide.* Columbia University Press.

Rajkumar, Andrew Sunil, and Vinaya Swaroop. 2008. "Public spending and outcomes: Does governance matter?" *Journal of development economics*, 86(1): 96–111.

Rao, Vijayendra, Indrani Gupta, Michael Lokshin, and Smarajit Jana. 2003. "Sex workers and the cost of safe sex: the compensating differential for condom use among Calcutta prostitutes." *Journal of Development Economics*, 71(2): 585–603.

Reinikka, Ritva, and Jakob Svensson. 2005. "Fighting corruption to improve schooling: Evidence from a newspaper campaign in Uganda." *Journal of the European Economic Association*, 3(2-3): 259–267.

Robinson, Jonathan, and Ethan Yeh. 2011. "Transactional sex as a response to risk in Western Kenya." *American Economic Journal: Applied Economics*, 3(1): 35–64.

Schouten, Erik J, Andreas Jahn, Anne Ben-Smith, Simon D Makombe, Anthony D Harries, Francis Aboagye-Nyame, and Frank Chimbwandira. 2011. "Antiretroviral drug supply challenges in the era of scaling up ART in Malawi." *Journal of the International AIDS Society*, 14(Suppl 1): S4.

Scott, James C. 1977. *The moral economy of the peasant: Rebellion and subsistence in Southeast Asia.* Yale University Press.

Sears, David O, and John B McConahay. 1973. *The politics of violence: The new urban blacks and the Watts riot.* Houghton Mifflin Boston.

Sen, Amartya. 1973. *On economic inequality.* Oxford: Clarendon Press.

Skaperdas, Stergios. 1992. "Cooperation, conflict, and power in the absence of property rights." *The American Economic Review*, 720–739.

Soares, Rodrigo R. 2004. "Development, crime and punishment: accounting for the international differences in crime rates." *Journal of Development Economics*, 73(1): 155–184.

Straus, Scott. 2006. *The order of genocide: Race, power, and war in Rwanda.* Cambridge Univ Press.

Svensson, Jakob. 2005. "Eight questions about corruption." *The Journal of Economic Perspectives*, 19(3): 19–42.

Thirumurthy, Harsha, Cristian Pop-Eleches, James Habyarimana, Markus Goldstein, and Joshua Graff Zivin. 2012. "Behavioral Responses of Patients in AIDS Treatment Programs: Sexual Behavior in Kenya." Vol. 15, 1–29.

Thirumurthy, Harsha, Joshua Graff Zivin, and Markus Goldstein. 2008. "The economic impact of AIDS treatment labor supply in Western Kenya." *Journal of Human Resources*, 43(3): 511–552.

Tilly, Charles. 1999. *Durable inequality.* Univ of California Press.

Verwimp, Philip. 2004. "Death and survival during the 1994 genocide in Rwanda." *Population Studies*, 58(2): 233–245.

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

Vian, Taryn. 2008. "Review of corruption in the health sector: theory, methods and interventions." *Health Policy and Planning*, 23(2): 83–94.

Wilson, Nicholas. 2011. "Antiretroviral Therapy and Demand for HIV Testing: Evidence from Zambia." *Available at SSRN 1982185*.

Wilson, Nicholas, Wentao Xiong, and Christine Mattson. 2012. "Is Sex Like Driving? Risk Compensation Associated with Male Circumcision in Kisumu, Kenya."

Winston, Clifford, Vikram Maheshri, and Fred Mannering. 2006. "An exploration of the offset hypothesis using disaggregate data: The case of airbags and antilock brakes." *Journal of Risk and Uncertainty*, 32(2): 83–99.

Yanagizawa, David. 2006. "Malthus in Rwanda? Scarcity, survival and causes of the genocide." *rapport nr.: Working Papers in Economics*, , (201).

Yanagizawa-Drott, David. 2010. "Propaganda and conflict: Theory and evidence from the rwandan genocide." Working Paper, Harvard University.

Young, Alwyn. 2005. "The gift of the dying: The tragedy of AIDS and the welfare of future African generations." *The Quarterly Journal of Economics*, 120(2): 423–466.

Young, Alwyn. 2007. "In sorrow to bring forth children: fertility amidst the plague of HIV." *Journal of economic growth*, 12(4): 283–327.