Microdynamics of War-to-Peace Transitions: Evidence from Burundi

Cyrus Samii

Submitted in partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2011

ABSTRACT

Microdynamics of War-to-Peace Transitions: Evidence from Burundi

Cyrus Samii

In these three essays, I study important facets of the transition to peace after Burundi's 1993-2005 civil war. The first essay studies the effects of quota-based ethnic integration within Burundi's army on expressions of prejudice and ethnic salience among soldiers. Exploiting a natural experiment within the army, I find that exposure to quota-based integration reduced prejudice and had no effect on ethnic salience, countering a prevailing view in the literature that quota-based integration is likely to exacerbate ethnic tensions. The second essay studies individuals' preferences over transitional justice alternatives. I find that support for punishment and truth-seeking is more tepid than the advocacy literature has suggested, that ethno-political motivations seem to dominate expressed preferences for punishment and truth-seeking, and, using a persuasion experiment, that simple forms of deliberation may actually polarize people. The third essay, co-authored with Michael Gilligan and Eric Mvukiyehe, examines the impact of Burundi's ex-combatant reintegration program on the economic and political reintegration of demobilized rebels. Exploiting another natural experiment, we find that the program provided substantial economic benefits, but that these economic benefits did not seem to contribute to political integration, at least in the short-run. The essays enrich our understanding of Burundi's difficult transition to peace. They also essays just how one may bring high scientific standards to study policies in the otherwise challenging context of post-conflict transitions.

Table of Contents

List of Figures

List of Tables

Acknowledgments

This work would not have been possible without generous mentoring and guidance from Macartan Humphreys, Michael Gilligan, Jack Snyder, and Andrew Gelman, and the terrific collaboration with my field research partners, Eric Mvukiyehe and Gwendolyn Taylor. The data used in this dissertation come from the multi-purpose *Wartime and Post-Conflict Experiences in Burundi Survey*, which received support from the Folke Bernadotte Academy, Sweden, the United States Institute of Peace, the New York University Department of Politics, and the Columbia University Graduate School of Arts and Sciences. I am grateful to my friends and colleagues at *Iteka – Ligue Burundaise des Droits de l'hommes* (Burundian League for Human Rights) for being able and supportive partners in our fieldwork. Christopher Blattman, Birger Heldt, Kimuli Kasara, Ronald Krebs, Kristin Michelitch, Suresh Naidu, Monika Nalepa, Philip Verwimp, and seminar participants at Columbia University, ISTEEBU (Bujumbura), MIT, New York University, UCSD, Uppsala University, and Yale University generously commented on component papers. I thank my parents for their support, and my biggest thanks go to my wife, Nicole, and my son, Caspar, for their patience and inspiration. All errors or omissions are my own.

For Nicole and Caspar

Chapter 1

Introduction

In these three essays, I study important facets of the transition to peace after Burundi's 1993-2005 civil war. Burundi is a small, impoverished, and land-locked country of approximately 8 million people (ca. 2010) in central Africa. Burundian politics is marked by two key structural features: (i) a caste-like division between an historically privileged minority group, the Tutsi, and an historically oppressed majority group, the Hutu, and (ii) clan-based and regional affiliations nested within the Tutsi-Hutu divide.[1] Since independence in 1962, Burundian society has endured a violent cycle of contestation organized around these cleavages. Protagonists in the conflict as well as commentators have characterized Burundi's bloody history in terms of two competing master narratives: on the one hand, a contest between certain Tutsi elite seeking to maintain extreme privilege and Hutus struggling for liberation from barriers to mobility, and on the other hand, and a contest between Tutsis seeking to protect their legitimately gained property and Hutus seeking to expropriate them under a genocidal rallying cry of purging "alien" Tutsi.[2] Another prominent perspective views

[1]In Kirundi, the national language, the appropriate manner to refer to the segment of society identified as Tutsi would be "Batutsi," and Hutu, "Bahutu." To keep the text lighter and to conform with conventions in Western texts, I simply use the shorter "Tutsi" and "Hutu."

[2]See, e.g., Lemarchand (1996) on the first perspective, and Chretien (2003) on the second.

the conflict in terms of elite manipulation in the service of zero-sum inter-regional and inter-clan resource competition under conditions of extreme scarcity (Ngaruko and Nkurunziza, 2000). The latest bout in Burundi's cycle of post-independence violence was a 1993-2005 civil war that touched all regions of the country and is estimated to have resulted in some 300,000 war-related deaths.[3] The essays in this dissertation examine in close detail the challenges that individuals faced, the strategies employed, and the medium-term results experienced in three crucial processes following the end of this latest civil war: (1) managing a painful legacy of inter-ethnic violence in constructing a new national army, (2) deliberating the appropriate manner for dealing with past human rights perpetrators, and (3) finding a way to reintegrate into post-war civilian life rebels who fought in a brutal, decade-long struggle. The essays rely on rich micro-level data collected a part of an original survey of over 3,000 Burundian adults, including current and former members of the army and rebel groups, as well as civilians who did not fight in the war.

My goals with these essays are two-fold. First, I seek to document in high resolution Burundi's societal reconstruction after the civil war. The reasons for doing so are not only to provide an historical record, but also to uncover potential lessons that might be applicable to cases elsewhere that face similar challenges. Indeed, the challenges of overcoming legacies of mass exclusion, inter-ethnic violence, and rebel reintegration are shared by many countries in transition.

Second, I seek to demonstrate how state-of-the-art quantitative social scientific methodology can be applied to addressing policy questions in this difficult research context. This research context is difficult not only because of the complexity of the

[3] I refer to 2005 as the end of this latest civil war. The date corresponds with the year of the general elections that consolidated the terms of the peace settlement between the incumbent regime and the major rebel factions. Nonetheless, I do recognize that splinter rebel factions have remained active in their fighting, even to this day. Nonetheless, relative to the violence from the previous decade, violence since 2005 has been much smaller scale and contained to very small enclaves around the capital city, Bujumbura.

issues, but also because of the dearth of readily-available administrative data as well as the risks to researchers and subjects alike associated with operating in an environment affected by violence. Nonetheless, by applying some ingenuity and patience, and by steeping oneself in the social and political nuances of the setting, I believe that it is possible to carry out policy studies that stand up to serious scientific scrutiny. In this way, I see myself as among the growing number of researchers that seek to work in the area of conflict studies while adhering to the strict principles of what quantitative methodologists, Angrist and Pischke, have labeled the "credibility revolution" in the social sciences (Angrist and Pischke, 2010).

A brief overview of the contents of these essays is as follows. The first essay studies the effects of quota-based ethnic integration within Burundi's army on expressions of prejudice and ethnic salience among soldiers. The essay exploits a natural experiment within the army, whereby exposure to quota-based ethnic integration in the army was based on an age cut-off. To measure prejudice, I use an experimental measurement technique, based on the random assignment of enumerators with different ethnicities. I find that exposure to quota-based integration reduced prejudice and had no effect on ethnic salience, countering a prevailing view in the literature that quota-based integration is likely to exacerbate ethnic tensions.

The second essay studies individuals' preferences over transitional justice alternatives. It studies why some individuals may prefer to "forgive" or "forget" rather than aggressively pursue punishment of offenders or truth about past abuses. Using a method that seeks minimize social desirability bias, I find that support for punishment and truth-seeking is more tepid than the advocacy literature has suggested. A correlational analysis shows that individuals tend to express preferences that are politically advantageous for their ethnic group, and that such apparent political motivations moderate the association between victimization and demands for punishment or truth. A persuasion experiment shows that simple forms of deliberation may actually polarize people. The results suggest the need for transitional justice practitioners to be cautious about upsetting hard-won political compromises.

The third essay, co-authored with Michael Gilligan and Eric Mvukiyehe, examines the impact of Burundi's ex-combatant reintegration program on the economic and political reintegration of demobilized rebels. We exploit a natural experiment within the program, whereby a contractual dispute resulted in about a third of beneficiaries from having their benefits withheld. Using these unfortunate ex-rebels as a pseudo control group, we find that the program provided substantial economic benefits, particularly to those who would otherwise have been among the worst off. These results stand in contrast to previous studies, however we note that these previous studies did not have an exogenous source of variation in program accessibility in the manner of our study. Nonetheless, we do not find that these economic benefits contributed to political integration, although our results can only speak to short-term effects.

Since the time that I began work on this dissertation, the community conducting quantitative micro-level research on conflict processes has grown from a small and tight-knit group to an expansive, multidisciplinary, and international research community. More and more high resolution data on violence and post-conflict processes comes online each month.[4] But data without a research design can say little. Findings from a research design that is not linked to well articulated theory cannot travel far. Finally, micro-level research designs and theories that are not well-adapted to the nuances of a local context can mislead. The challenge for the "microdynamics of conflict" research program is to apply the principles of the "credibility revolution" for robust causal empirical analysis, test well articulated and generalizable theory, and operationalize such theory in a way that is sensitive to nuances of the local context. My hope is that the essays contained in this dissertation are a step in that direction.

[4] See Blattman and Miguel (2010) for a relevant review.

References

Angrist JD, Pischke JS. 2010. "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics." *Journal of Economic Perspectives*. 24(2):3-30.

Blattman C, Miguel E. 2010. "Civil war." *Journal of Economic Literature*. 48(1):3-57.

Chretien JP. 2003. *The Great Lakes of Africa: Two Thousand Years of History, translated by Scott Straus*. New York: Zone Books.

Lemarchand R. 1996. *Burundi: Ethnic Conflict and Genocide*. Cambridge: Cambridge University Press.

Ngaruko F, Nkurunziza JD. 2000. "An economic interpretation of conflict in Burundi." *Journal of African Economies*. 9(3):370-409.

Chapter 2

Do quotas exacerbate or reduce ethnic conflict? Micro-level evidence from Burundi's military

Summary

An important issue in political development is why conflict between ethnic groups endures and how such conflict might be transcended. Using original survey data from post-war Burundi, I examine whether quota-based integration of ethnic groups in the military exacerbated or reduced the kinds of prejudice and ethnic salience necessary for ethnic conflict to endure. Among members of the incumbent armed forces, eligibility for the new integrated armed forces was based on an age cut-off. By comparing those just above and below the cut-off, I isolate the effects of being exposed to involuntary ethnic integration in the military versus being demobilized. Then, using tests for changes in other variables at the cut point, I examine what these effects suggest about the effects of exposure to ethnic integration per se. I find that participation in the quota-integrated military is associated with decreases in prejudicial behavior and appears to be benign in terms of effects on ethnic salience. I find that these results are not likely due to adverse shocks associated with being

demobilized.

The results have important implications for policy. They suggest that pessimism about quota-based integration, based on a presumption that they intensify or "freeze" conflicting ethnic identities, may be unjustified. In Burundi, Hutu and Tutsi identities have divided society for generations and have been central in a decades-long cycle of violence. The evidence in this paper weighs slightly in favor of quota-based integration as a strategy for overcoming legacies of exclusion and inter-group conflict.

2.1 Introduction

An important issue in political development is why conflict between ethnic groups endures, and how such inter-group conflict might be transcended. Quota-based ethnic integration is a commonly used policy, based on the apparent necessity of using proactive measures to overcome legacies of ethnic exclusion. However, detractors have proposed that such quotas may exacerbate the very conflict they are intended to transcend. I use original survey data collected in Burundi after the end of their 1993-2004 civil war to study the effects of exposure to quota-based ethnic integration in the national military on soldiers' expressions of ethnic prejudice and ethnic salience. The analysis allows me to test competing claims about whether exposure to such quotas reduce or intensify the kind of behaviors and perceptions that contribute to the endurance of ethnic conflict. I use a research design that takes advantage of a natural experiment on exposure to quota-based integration as part of participation in Burundi's reformed military.

Hutu and Tutsi identities have divided Burundian society for generations, providing the master cleavage for a cycle of massive violence since independence. At the heart of this violence has been the problem of an "ethnicized" military. The rebellion initiated in 1993-1994 was led by Hutu soldiers who defected from the military in order to destroy what they labeled a "mono-ethnic," Tutsi-dominated army. The 1993-2004 civil war resulted in a peace agreement that redistributed power and called for the ethnic integration of the military. The agreement established a 50-50 Hutu-Tutsi ethnic quota. Among members of the former national military, an age-based cut-off (below 45 years of age in 2006) was used to determine eligibility to serve in the new national military. By comparing those just above and below the cut-off, I isolate the effects of being exposed to ethnic integration in the military versus being demobilized. I examine the extent to which my estimates may be tainted by other differences between the experience of participation in the integrated military versus being demobilized. The results show that participation in the quota-integrated military is associated with less prejudicial behavior and appears to be benign in terms of its effects on ethnic salience. It does not seem that these differences are due to shocks associated with demobilization. Thus, the evidence suggests that exposure to ethnic integration may reduce tendencies necessary for ethnic conflict to endure, and at worse do little to exacerbate them.

There are a number of reasons that levels of ethnic harmony or disharmony within the military are crucial in political development. Accounts of the erosion of political order and the outbreak of mass conflict often point to "ethnicization" within the military as a fundamental factor, for example in the onset of the Biafra War in Nigeria (Luckham, 1971) and the First Liberian Civil War (Adebajo, 2002). The responsiveness of military members to ethnic appeals by politicians is affected, at least at the margins, by the degree of solidarity across ethnic lines within the army. Ethnic relations among members of the military affect the ease with which "subjective control" of sections of the military might be achieved by political entrepreneurs who use ethnic appeals (Huntington, 1957). Constraining ethnic factionalization in the military is especially important in developing countries, particularly countries in Africa. In these contexts, military intervention and political manipulation of the military along ethnic lines has been a central feature of the dynamics of political order and disorder (Horowitz, 1985; Hutchful, 1997). Given its dominance in the institutional landscape, militaries in developing countries have enormous symbolic value. Therefore, levels of harmony within the military are likely to influence mass attitudes about the potential for inter-ethnic cooperation. In Burundi, this has certainly been the case. A history of ethnically-colored army repression provided a focal point for genocidal mass mobilization in response to a bungled 1993 coup (United Nations, 1996). The ethnically charged interpretation of the army's actions facilitated in no small part the organization of genocide in neighboring Rwanda some months later.[1] With respect to political order, inter-ethnic relations among members of the military are among

[1] See Samii (2010a) for a detailed discussion.

the most important in society.[2]

This study contributes to political scientists' understanding of how policies enacted as part of peace processes may *in themselves* contribute to either the perpetuation or transcendence of ethnic conflict. Ethnic quotas are common features of peace processes, although little research has been done to evaluate their effects. If we find that quotas contribute to entrenching or exacerbating conflict, then we would want to consider alternatives or ways to complement quotas with policies to mitigate these effects. If the effects are to reduce prejudice, then the recommendation would be to highlight these effects, setting an example that might inspire easing of ethnic tension more generally in society. Some may argue that such "optimal design principles" are irrelevant for peace negotiations, based on the presumption that if lines of conflict are preserved in a peace agreement, then this is probably intentional as a way for one or another party to maintain the credibility of a threat to resume fighting should their counterparts renege. This view suggests there can be little scope for the independent tinkering to design transitions that reduce the scope for conflict. I think this view is misrepresents what happens in peace processes. External mediators exert considerable influence on the shape of peace agreements. Parties to a dispute often approach mediators searching for suggestions on institutional arrangements that can help to translate a ceasefire into a more durable peace. In the case of Burundi that I study here, South African mediators and technical experts provided many ideas and greatly influenced the institutional reform process.[3]

---

[2]Krebs (2004) takes a mostly contrary view, proposing military integration is unlikely to contribute to reducing inter-group conflict in society at large. But his argument is based mostly on a *lack* of solid evidence on the effects of socialization, rather than clear evidence pointing to no effect. This paper is attempting to provide more credible evidence. In addition, Krebs acknowledges that if military members have influence in politics, then socialization from within the military may be societally important. In Burundi and in cases like Burundi, this condition certainly holds.

[3]I provide a more detailed account of the military integration process, highlighting South Africa's role, in Samii (2010a).

The context of this study is contemporary Burundi. There are two reasons that Burundi is an important case for students of current transitions from war to peace. First, structurally, Burundi's Hutu-Tutsi ethnic structure is part of the class of "ranked ethnic systems" that have provided the setting for violent conflicts that have been especially difficult to untangle (Horowitz, 1985; Wimmer, 2006). The privilege afforded to segments of the minority Tutsi segment and the severe constraints to mobility for majority Hutu provide a ready narrative for ethnic mobilization that has colored Burundi's, and neighboring Rwanda's, violent post-independence history. Lessons from this case may be applicable to other contexts where the we face the challenge of transcending the legacy of mass exclusion and perceived minority ethnic domination. Second, Burundi's location in Africa's Great Lakes has unfortunately been host to an enormous fraction of war-related death in recent decades. In addition to Burundi's fragile peace, there continues to be much worry about the stability of Rwanda's post-war political order. The nearly decade-and-a-half of violent disorder in the east of neighboring Democratic Republic of of Congo is partially a result of spill-overs from Rwanda and Burundi. Understanding the nature of contemporary civil conflict is, to no small extent, a matter of understanding this particular set of intertwined conflicts.

This paper begins with a discussion of hypotheses on the effects of quotas, focusing on schools of thought associated with the contact hypothesis and various theories of hierarchy maintenance. I follow with a discussion of the context of post-war military integration in Burundi, explaining how the conditions provide for an interesting test of these hypotheses. I then present the methods used in the study, explaining the sample survey and regression discontinuity design used to identity causal effects. I follow with the data analysis and results, as well as a discussion of robustness of the findings. A conclusion summarizes.

2.2 Quotas and ethnic conflict

There exists a theoretical debate in the social science literature about the consequences of quotas for ethnic conflict. Quota-based integration of institutions is a commonly employed strategy in addressing legacies of ethnic exclusion and ethnic conflict, due in part to their transparency and to the appreciation that legacies of exclusion may not disappear without proactive measures (see, e.g., Fryer and Loury, 2005). Aside from the philosophical debate over quotas, the behavioral science literature features debate over the likely consequences of quotas. This includes an emerging literature on the effects of quota-based representation in elected bodies, where results suggest that quotas and other types of "descriptive representation" may contribute to enhanced attention to marginalized groups' welfare (Duflo, 2005) and less prejudicial attitudes and behaviors among constituents (Dunning, 2010; Chauchard, 2010). I focus here on theories about how individuals subject to quota-integrated institutions may react and the implications of such reactions for ethnic conflict more generally. We can distinguish between those who are optimistic about the contribution of quotas to reducing ethnic tensions from those who are pessimistic, suggesting instead that quotas may exacerbate tensions.

Optimistic theories derive mostly from the theoretical tradition tied to the intergroup "contact hypothesis" (Allport, 1954; Pettigrew, 1998). The contact hypothesis suggests that increased contact in the presence of four enabling conditions will lead to a reduction in negative affect toward an out-group and a consequent reduction in prejudice and parochialism. These enabling conditions, as applied to a quota-integrated institution, are that (1) the integrated groups are afforded equal status in the institution, (2) members of the institution work toward common goals, (3) attainment of those goals must require cooperation, and (4) integrative aims must have the support of authorities, law, or custom. The supposed psychological mechanisms through which the reduction in negative affect occurs are that subjects correct misheld views about out-group members by learning about each other, develop a sense that working with out-group members is normal, form emotional bonds to individuals across group boundaries, and learn that in-group norms are not necessarily optimal or even superior to those of the out-group. In a meta-analysis of 515 studies, Pettigrew and Tropp (2006) find that the hypothesis holds in a broad array of contexts. Political scientists have applied contact theory in a number of settings, including in analyzing the possible effects of integrative policies in post-apartheid South Africa (Gibson and Gouws, 2003) and in interpreting how a reconciliatory media intervention helped to reduce inter-group prejudices in post-war Rwanda (Paluck and Green, 2009). Ethnic integration in the military provides a good test of these hypotheses, as the formal terms of integration typically establish enabling conditions (1) and (4), and the nature of military activity establishes enabling conditions (2) and (3). Thus, by working together as "brothers in arms" in an integrated military, contact theory leads us to expect that soldiers would become less ethnically prejudiced and less parochial.

Other theoretical approaches suggest reasons to be more pessimistic. In the social psychology literature, theories of hierarchy maintenance and status preservation suggest that quota-based integration will generate resentment among those whose status is threatened, causing such individuals to withdraw cooperation, become more adversarial, and even work to undermine the quota-integrated institution (Blumer, 1958; Coser, 1956; Levine and Campbell, 1972; Spilerman, 1970). In the political science literature, hypotheses along these lines have been used to explain the origins of violent inter-group conflict, sometimes pointing specifically to status shifts within major state institutions such as the military as precipitating causes (Gurr, 1970; Horowitz, 1985, 2001; Petersen, 2002). In a similar vein, theorists of civil conflict have proposed that the use of quotas may freeze the salience of ethnic identities that political entrepreneurs had used instrumentally during wartime to divide people and seek political advantage (Aitken, 2002; Simonsen, 2005). If this is true and if soldiers are really the muscle of ethnicized political tendencies, then we should expect soldiers to be especially adamant about the importance of their ethnic identity.

Each set of hypotheses is plausible, and so this study aims to assess the relative strength of integrative versus oppositional effects of quota-based ethnic integration. The overall effect of quota-based integration is likely an aggregate of both positive and negative effects. At the end of the day, our interest for policy is to determine whether one or another tendency dominates.

2.3 Context

The context is the ethnic integration of Burundi's military after the 1993-2004 civil war. Integration of mostly Hutu rebels into the Tutsi dominated army began after a 2003-4 ceasefire which drew into the peace process the largest rebel group in the country, the National Council for the Defense of Democracy-Forces for the Defense of Democracy (CNDD-FDD, by its French acronym).[4] Some historical background is necessary to understand why ethnic integration was such a landmark step in Burundi's political development. Burundi is a small, impoverished, land-locked country of approximately 8 million people (as of 2010) in the center of Africa. It has been wracked by a cycle of political violence since independence in 1962. Like neighboring Rwanda to the north, Burundian society is marked by a caste-like stratification that has historically privileged a Tutsi minority relative to majority Hutu and a very small third group, the Twa.[5] Also like their neighbours in Rwanda, Burundians have strug-

[4]This section provides only a brief overview of the context. A detailed account, based largely on my own interviews with Burundian military and political leaders as well as technical advisors to the process, is given in Samii (2010a).

[5]Conventionally, Tutsi are said to constitute 14% of Burundian society, Hutu 85%, and Twa 1%. These figures are from a 1956 colonial-era census of dubious quality. The current distribution is likely to differ, not least due to imbalances in mortality rates in the various crises since independence. Analysis of survey data collected by my research team in 2007 suggests that the distribution may slightly overstate the Hutu proportion, although the margins of error are quite large. Nonetheles, to the extent that electoral results from 1993 and 2005 largely reflect ethnic preferences, the 14-85-1

gled to escape a conflict pitting custodians of this "ranked ethnic system" (Horowitz 1985; Lemarchand, 1970) against those ostensibly seeking to remove barriers to Hutu mobility.

Burundi's national army, known as the *Forces armées burundaises* (FAB) until 2004, has featured centrally in the country's bloody political drama since independence in 1962. In the first four years after independence Burundian politics suffered a series of assassinations, an abortive coup by Hutu officers, repressions, and reprisal massacres. The events culminated in a purge of high ranking Hutu officers and a 1966 military coup led by the minister of defense, Captain Michel Micombero, a Tutsi from Bururi province. Thus began a period of *de facto* military rule and intensified concentration of economic opportunities and power, particularly within the army, into the hands of a Tutsi clique from the southern Bururi province. The clique oversaw a dramatic intensification of Hutu exclusion as well as a degree of exclusion of non-southern Tutsi. A 1972 insurrection coordinated by Hutu expatriates and Hutu army members escalated to involve massacres of Tutsis, mostly in the southern part of the country. This triggered a barbarous crackdown by the army, which went beyond restoring order and sought to prevent future uprisings by "decapitating" Hutu society. A more thorough purge of Hutu members of the army and police followed. Competition among clan-based factions of southern Tutsi military officers shaped the next 20 years of Burundian politics.

In the late 1980s and early 1990s, then-President Pierre Buyoya, also a Tutsi from Bururi, presided over a process of ostensible national reconciliation. Buyoya oversaw the promulgation of a national unity charter and new constitution in 1992, setting the stage for elections in 1993. Some places in the national officer academy, the *Institut supérieur des cadres militaires* (ISCAM), were opened to Hutu candidates. But the gesture masked a more general resistance to army reform among the Tutsi elite. The 1992 national unity charter declared that "the truth is that there is no discrimination

distribution may not be so far off.

within the army," a rather absurd claim (Lemarchand, 1996:139). One of the beneficiaries of this process, a Hutu from Bururi named Jean-Bosco Ndayikengurukiye, was a member of one of the integrated ISCAM classes. He would eventually defect to become a leader in the rebellion that broke out in 1993.

In peaceful and generally fair elections in 1993, Melchior Ndadaye, a civilian Hutu who had returned from exile in Rwanda, defeated the incumbent Buyoya by a large majority in the presidential race. Ndadaye's administration called for the rapid promotion of some Hutu officers within the military, ostensibly to better align the military officer corps with the interests of the civilian government. After only 3 months in power, Ndadaye was assassinated in a bungled coup attempt on October 21, 1993. The assassination triggered what a United Nations commission described as genocidal reprisals by Hutu mobs against Tutsi men throughout the countryside, followed by massacres of Hutus by the mostly Tutsi army and police (United Nations, 1996). Hutu members of Ndadaye's government fled the capital, Bujumbura, to establish a rebel movement, the CNDD, and its military wing, the FDD. Explicit in their stated goals was the defeat and dismantling of the so-called "*armée mono-ethnique*," so-called because the officer corps was the near-exclusive domain of southern Tutsis.

The army (FAB), CNDD-FDD, and some smaller rebel factions fought in a civil war that lasted until 2004. Fighting touched most regions. It resulted in deaths estimated at about 300,000 out of a population of 6-8 million. Burundi's pre-war socio-economic development levels were already among the world's lowest, although for its income level, the country did have relatively well-developed infrastructure and institutions. The war severely stalled development for over a decade, resulting for example in an estimated 20% decline in real GDP over 1993-2002 (World Bank, 2004).

A peace process had begun in 1996 and discussions of military integration featured prominently from the start. Agreements signed by the warring parties in Arusha in 2000 and Pretoria in 2003 ushered in genuine peace. The FDD forces were largely successful on the battlefield, although the FAB forces were not defeated outright. These rebel successes are reflected in the agreements, whose provisions constitute a near revolution in the country's distribution of power. Key among them was integration of the army and police. The peace agreement established a rule of "ethnic balance" such that posts would be allocated to Hutus and Tutsis in a 50-50 manner, and the overall composition of the security forces was to be balanced in this way "in view of the need to achieve ethnic balance and to prevent acts of genocide and coups d'état" (*Arusha Accords*, Protocol II, Chapter 1, Article 11). The agreement called for the integration of members of the rebel groups into a reconstituted national military. Technical experts from South Africa provided substantial input into the process, drawing on their own integration experience.[6] Precise details of integration were finally set in a "Forces Technical Agreement" signed in Pretoria in November 2003. It called for an integrated army top officer echelon with 60% FAB officers and 40% CNDD-FDD officers, and a 65-35 FAB to CNDD-FDD breakdown for the integrated police top officer echelon. Throughout the ranks of these combined security forces, the principle of "ethnic equilibrium (50/50)" would be observed.

A two-phase integration process was set in motion in 2004. The first phase was the "assembly" phase, which ran from 2004 to late 2005. During this phase, some 26,000 members of the rebel armies were gathered in cantonments, with a contingent of 7,000 CNDD-FDD soldiers immediately merged into integrated units with members from the 40,000-strong FAB. During this period, 14,000 soldiers were demobilized, of which about 5,000 came from the FAB and 9,000 from the rebel forces. The second phase, beginning in late 2005 was the "rationalization" phase. It was during this phase that full integration took place, including bridging training for rebel combatants and full mixing of ex-national army members and the remaining 10,000 ex-rebels into common units. During this time, the military was also trimmed toward a target of 25,000

[6] The peace process was ushered forward by sustained intervention by no less than Julius Nyerere, Nelson Mandela, Bill Clinton, and Thabo Mbeki.

army troops and 20,000 police. Among FAB troops, the first wave of demobilization in 2004 was mostly voluntary, as the military was too pre-occupied with the task of "assembly" to process retirement. Then, the 45-year age cut-off was applied in with regularity only in the later phases, starting in late 2005.

[Figure 2.1 here.]

The mixed units trained together, stayed in the same barracks, and fought together in both counter-insurgency operations against factions that remained active as well as in contributions to peacekeeping missions in Sudan and Somalia.[7] Figure 2.1 shows the sample proportions of Hutu and Tutsi military members from each of the barracks that from which samples were drawn for this study (the sample is described in detail below). We see that the ethnic integration called-for in the peace agreement was apparent across the barracks at the time of the field research. This integration process was a highly salient feature of the post-war transition in Burundi. Indeed, in a survey of Burundi civilians conducted in 2007 in parallel to this study of military integration, 63% of respondents listed military integration as among the main points of the peace agreement, despite not having been prompted to do so.[8]

Ethnic integration and the resulting forced inter-ethnic contact are the dominant thing that distinguishes the post-late-2005 military from that which preceded it. As discussed in Samii (2010a), the Forces Technical Agreement that governed the creation of the new military called for the retention of the old FAB rank structure training protocols, and formal doctrine. Conflict resolution training programs were issued to members of the new army, but these were also delivered to demobilized soldiers. A special reconciliation program, the Burundi Leadership Training Program (Wolpe and McDonald, 2006), was limited to upper echelon officers. This study focuses on rank-and-file members of the military, and so I am doubtful that any effects that we

[7]See Samii (2010a) for more details.

[8]These data are analyzed in Samii (2010b).

measure here can be attributed to those efforts.

2.4 Methods

2.4.1 Sample

The data on soldiers and are drawn from the multi-purpose *Wartime and Post-Conflict Experiences in Burundi* survey.[9] The survey includes data from interviews with civilians, as well as ex-rebels and ex-army, both demobilized and those integrated into the new security forces. This paper works with only data on men who were professional rank-and-file members or non-commissioned officers in the national army—the *Forces armées burundaises* (FAB)—during the war.[10] Substantively, this is not such a major compromise, as it is FAB members who offer the most interesting test of the hypotheses given above. It is for them that the integration process spelled a major loss in relative standing. Therefore, it is for them that we can genuinely assess the relative strength of "resentment" versus "contact."

The sampling strategies were slightly different for demobilized FAB versus those in the integrated military. Demobilized FAB were selected through a multistage random sample from lists of demobilized ex-combatants registered to receive reintegration benefits through Burundi's national demobilization, disarmament, and reintegration (DDR) program. The first stage involved randomly selecting 66 out of Burundi's 129 *communes*—Burundi's second-tier administrative unit—from a list of communes stratified by Burundi's 17 provinces. We then set a target number of ex-combatant interviews to complete in each commune, with targets proportional to the national

[9]Details on the survey are available at www.columbia.edu/~cds81/burundisurvey/

[10]It is only for them that the age-based cut-off that I exploit was binding to any significant degree. While I would like to study outcomes associated with members of the former rebel forces, doing so would require identifying some other source of plausibly exogenous variation in participation in the integrated armed forces. Not having that, I limit the scope to FAB members.

proportion of ex-combatants registered with the DDR program in the commune. Targets ranged from 2 to 33. Then, we obtained from the national DDR office the complete lists of ex-combatants registered as residents of each of the selected communes. We then drew a simple random sample (with a random number generator in the software package, R) of the desired number of interviews from each of these commune lists. We created a randomly selected reserve list for each commune to use in the case of non-response. Selected participants were contacted and brought to the respective Provincial Bureau by DDR program staff for interview on a scheduled date by our enumeration team. Active members of the military were selected through a separate multistage process. In the same 66 communes as described above, enumerators approached police station chiefs with a letter from the Ministry of Interior explaining that they were to list all officers stationed in the commune. Then the enumerators randomly selected a target number of police which was proportional to the population size of the commune. The reason is that police assignments are given in a manner proportional to commune population. This creates as close to a self-weighting, representative sample as possible with available information. For the active members of the military, we identified all camps listed in the 66 selected communes. Enumerators then approached the barrack with a letter from the Minister of Defense instructing them to cooperate with the enumeration team to use interval sampling to select non-commissioned officer and rank-and-file soldiers from the camp's register, with the number selected in each camp proportional to the total in the camp, as inferred from information provided by the Defense Ministry. The rate at which first choices were interviewed in each of these samples was very high—around 90% for each—and so I assume no need for further adjustment to account for non-response.[11]

[11]This very high response rate was likely due to a few factors: (1) respondents probably took the interview to be a requirement given the formal manner in which they were approached; (2) we accommodated schedules by setting dates for interviews well in advance; (3) the fieldwork was conducted during the idle interim period between planting and harvesting seasons, and so respondents working in the agricultural sector faced few competing demands on their time; and (4) for the de-

One can get a sense of the general relevance of the results presented in this paper by considering the location of the sample members in Burundian society. As discussed in the previous section the army was dominated by a Southern, Tutsi officer corps. The rank-and file included some Hutus, many of whom joined formally for sustainable employment after they completed a military track for Burundi's mandatory two-year public service requirement. However, the ethnic distribution is dominated by Tutsis: 892 out of 1086 (82%) of the FAB soldiers in our sample are Tutsi. Entrance into the rank and file, both in peacetime and during the war, was obtained as part of recruitment calls that were issued at least once a year—and sometimes more frequently during periods of more intense fighting during the war. To gain admission, one needed to pass an exam that included written components. Given the level of illiteracy in the country, this meant that even the rank-and-file excluded the ultra-poor segments of society. At the same time, because of the relative hardship of military life, few in the upper socio-economic strata would likely join the army, meaning that most members were drawn from a middle socio-economic segment.

2.4.2 Exogenous variation in exposure to ethnic integration in the new military

[Figure 2.2 here.]

Figure 2.2 shows survey respondents by military/demobilization status and distance from the age eligibility cut-off. The "treatment" here is participation in the integrated military and the "referent" or "control" condition is being demobilized and thus not serving in the integrated military. I appreciate that there are factors

mobilized soliders, while participation was voluntary, a "transport allowance" of about 2 US dollars was provided to each respondent after they completed the interview, thus making it worthwhile for respondents to sit through the entire interview. (Enumerators purchases cases of soft drinks for police stations and army camps.)

associated with demobilization status that mean that the comparison to those in the integrated military is an imperfect way to test hypotheses about quota-based integration per se. I address strategies for isolating the effects of exposure to the quota below.

At the bottom of Figure 2.2, I display a jittered rug plot of members of the sample participating in the intergration military over an eligibility-cutpoint-centered age variable. At the top of Figure 2.2 is a jittered rug plot of demobilized soldiers over the age variable, broken down into those who self-reported that they voluntarily demobilized (Demob vol.) versus those who self-reported that their demobilization was not voluntary (Demob invol.). Also plotted in Figure 2.2, with the gray dots, is the proportion of sample members, by age, that had been demobilized, computed within 5-year bins whose boundaries are marked on the bottom axis. We see that the demobilized proportion drops smoothly until the point marked 0, where it then shoots up discontinuously. Using a linear regression estimated within a 5-year window on either side of the cut-point, the estimated magnitude of the jump at the cut-point is 0.58 (robust s.e., 0.11).[12] This jump at the point marked zero provides the basis for identifying causal effects of being in the integrated military versus being demobilized. It is evident from the graph that no other such jumps exist in the demobilized proportion over the range of the age variable.

At the time of fieldwork in June-August 2007, eligibility for participation was restricted to those below the age 45 as of September 2006, the timing of the last application of the eligibility threshold for demobilization prior to fieldwork. Those who were less than 45 years of age had not yet been subject to the eligibility cut-off from the previous year. Those aged 45 and in the military would be subject to the eligibility criterion sometime shortly after field work as part of the rationalization process. Some of these people would have turned 45 in the months between September 2006 and the field work period of June-August 2007, and so the maximum age among

[12] Using a triangular kernel inside the same window, the estimate is 0.48 (robust s.e., 0.13).

those deemed eligible by the rules should be 45, and the youngest age for which there should be no ambiguity of ineligibility is 46, with some indeterminacy about ineligibility for those aged 45. I set the cut-off at 45.5 years to account for the coarseness. Looking at the data, we see that this results in some fuzziness at the cut-point. Focusing for a moment on fuzziness above the cutpoint, part of this is due to the type of timing issues for 45-year-olds mentioned. But some of it cannot be explained that way, and must be due to some idiosyncratic application of the eligibility cut-off. However, the numbers are very few.

Individuals in the military at the time of fieldwork would have been subject to about 18 months of experience in the integrated army. As discussed above, the eligibility threshold was only applied with regularity during the transition as of late 2005, just at the onset of integration. Our sample includes individuals who were demobilized as part of a wave of applications of the eligibility threshold in October-November 2005 (accounting for 55 out of the 85 retirements in the sample), with a second, smaller wave occurring in March 2006, and sporadic retirements otherwise. Thus, a few of the 45 and 46 year-old demobilized soldiers in the sample had the chance to participate in the integrated army for some period up to mid/late 2006, implying that they would have had some 6-9 months of exposure, although this would have ended at least 6 months prior to field work. Nonetheless, to make the estimation more straightforward, I consider these to be on the "non-treated" side of the eligibility cut-off. This coding will dilute the effect estimates to a certain extent, but the payoff is to make the estimation much more straightforward.

Potentially a larger concern is that anyone could either volunteer to demobilize or be selected for demobilization due to judgments of being "unfit to serve," irrespective of age. Interestingly, however, we see that the rate of demobilization decreases smoothly in age until the cut-point. As indicated as well on the graph, the vast majority of younger demobilized soldiers chose to exit the military voluntarily. The likely reason for the pattern is that voluntary demobilization would have been more attractive for younger combatants, given that they may have seen themselves as having more chances in starting a life outside the military. For this reason, voluntary demobilization would become less appealing as one grows older, with only forced retirement being the way to get an older person out.[13] This is apparent on the graph, as the ratio of voluntarily demobilized approaches zero as we reach the cut-point from the left. On the other side of the cut-point, we do see that the proportion of voluntarily demobilized increases a bit. Based on interviews with demobilization program staff, I understand that this was due to the temporal coarseness with which demobilization waves were carried out. The national demobilization program made opportunities to demobilize available at certain discrete moments, based on practical reasons. Soldiers who had not yet been forcibly retired but were approaching the age for this to happen could elect to demobilization during these opportunities, appreciating that they would still qualify for pension benefits. Because of the sparseness of the number of demobilized to the left of the cut-point, there is no way to tell whether the rate of voluntary demobilization to the right of the cut-point is unusual. But because they represent a minority of cases (about 24% when looking within the 5-year bin above the cut-point) and because there is a good explanation for it, there does not seem to be reason for concern about sorting on this basis around the cut-point.

2.4.3 Identification of the effects of participation versus demobilization

My research design allows me to identifying the causal effects of participation in the integrated military relative to demobilization. Below, I discuss how I use auxiliary analyses to try to isolate the effects of exposure to quota-based integration per se.

The research design is a "fuzzy" regression discontinuity design (Imbens and

[13] In Gilligan et al (2010), my co-authors and I study impacts of demobilization and reintegration program assistance in detail.

Lemieux, 2007). A fuzzy regression discontinuity design is a generalization of the "sharp" regression discontinuity design. The sharp design is applicable when a cut-off on a numerical index, called the "forcing variable," deterministically establishes what units are in or out of treatment. Examples of sharp designs in the political science literature include the use of electoral margins to determine which candidate in a two-candidate election is promoted to office. In those examples, the treatment is promotion to office, being assigned deterministically by whether one's vote share minus one's opponent's vote share is greater than 0 (Lee, 2008; Hainmueller and Eggers, 2009). A fuzzy design is a generalization of this, in that treatment assignment is not deterministically set by the position relative to the cut-off. Rather, treatment assignment is probabilistic over values of the forcing variable, but exhibits a substantial "jump" at the value of the cut-off. Such is precisely what we can see in Figure 2.2.

Given such a discontinuous jump in the treatment assignment probability, we can identify the effect of the treatment, for people defined by the value of the forcing variable at the cut-point, so long as a few assumptions hold.[14] First, expected values of the outcomes of interest for the treated, on the one hand, and the controls, on the other, are smooth in the immediate vicinity of the cut-point. Second, in the immediate vicinity of the cut-point, treatment status is unconfounded relative to the outcomes of interest, conditional on the forcing variable.[15] When these conditions hold, we can use an instrumental variable strategy to estimate the average treatment effect for individuals defined by the value of the forcing variable at the cut-point. The excluded instrument is an indicator variable for whether the person is above or below the cut-point. The treatment, in our case participation in the integrated military, is

[14] For formal statements of these assumptions, refer to Hahn et al (2001) and Imbens and Lemieux (2007).

[15] The second assumption is necessary for nonparametric identification with heterogenous treatment effects. See Hahn et al (2001), Theorem 2.

treated as an endogenous regressor. Treatment effects can then be estimated with two-stage least squares, with robust standard errors providing the appropriate basis for inference (Imbens and Lemieux, 2007:15-17).[16]

Given these identifying assumptions, the analyst must decide on a way to model the smooth relationship between the forcing variable and outcomes and must also decide on the size of the window within which to fit the model to measure effects. Imbens and Kalyanaraman (2009) have proposed a purely data-driven method to select an "optimal" bandwidth to use for using local linear regression. The method relies on asymptotic properties, which may not be well-motivated given the moderate sample size with which I am working. An alternative approach is to pick a window based on substantive knowledge about what ranges of differences in the value of the forcing variable are meaningful. In a sufficiently small window, one may reasonably claim local linearity in the relationship between the forcing variable and the outcomes and fit a linear model. With a wider window, a common strategy is to fit high order polynomials, reducing the order of the polynomial on the basis of statistical significance tests on higher order coefficients. Green et al (2009) studied the performance of these methods in recovering an experimental benchmark. They find that local linear regression with optimal bandwidth selection, linear regression within a small substantively chosen window, and the polynomial specification search strategy in larger windows all performed rather well, although their tests used a very large dataset (over 13,000 observations).[17] Based on the outcome measures that I use (discussed below), the

[16]For further robustness, I compute standard errors that account for clustering at the level of the *commune*, which was the primary sampling unit above that of the individual respondents.

[17]Imbens and Lemieux (2007) also propose a "cross-validation" based technique for selecting a window. However the technique requires the rather arbitrary choice of a maximum window size, Green et al (2009) do not find that this method delivers any gains over either the data driven "optimal" window selection or substantive window selection, and Imbens and Kalyanaraman (2009) propose that the "optimal" method should dominate over cross-validation. Thus, I do not employ the cross-validation technique.

Imbens-Kalyanaraman selects 3 years on either side of the cut-point as the optimal window.[18] This leaves me with a very small sample size (65 observations below and 45 above the cutpoint). Based on substantive considerations, a window of 5 years ought also to be tight enough to invoke local linearity, and this allows me to work with much more ample sample size (92 observations below and 68 observations above). I present results in both of these windows. To check whether finite sample size variability is affecting the analysis, I also present results for a wider window (10 years), using the polynomial specification search strategy. Imbens and Lemieux (2007:11) argue that given a sufficiently tight window, a local linear regression with a rectangular kernel (that is, an ordinary linear regression within the window) should be adequate. Some modicum of bias reduction may be gained with a more sophisticated kernel—e.g. a triangular kernel, equivalent to using weights that decrease linearly in distance from the cut-point—although at some cost to efficiency. Because efficiency concerns predominate in my analysis, I present results for the rectangular kernel, indicating in the text how things differ when a triangular kernel is employed.[19]

Regression discontinuity designs are heralded for their high internal validity relative to other observational study methods. Aside from the two assumptions given above, no further assumptions of unconfoundedness are needed. This means that the analyst does not need to search for a covariate set or specify a multivariate model. A limitation is in external validity. Effects are identified only for those units defined by the value of the forcing variable at the cut-point. For the purposes of this study, this would mean for former-FAB soldiers aged between 45 and 46 years of age. While this is a limitation, I think it also provides estimates for an especially interesting subgroup in the context of Burundian politics. It is these individuals who, because of their age and experience, provide important role models to their neighbors and peers. We should also expect them to be relative "set in their ways," and so this

[18]The method is implemented in Stata with a script made available by Imbens on his website.

[19]Results from the analysis with the triangular kernel are available from the author.

group provides a rather hard test for theories of attitude change.

2.4.4 Isolating effects of exposure to the quota

The discontinuity ensures that exposure to the quota was not due to self-selection. This is hugely important. Self-selection into or out-of the quota-integrated army would likely be a large source bias in using a comparison of integrated versus demobilized army members to estimate the effect of exposure to the quota. Presumably, those who are more prejudiced or for whom ethnicity is of high salience would choose not to participate, in which case such a simple comparison would overstate the effect of exposure to the quota. The discontinuity that I exploit removes that important source of potential bias.

But the regression discontinuity only narrows things down to a certain extent. Other contextual factors may differ on either side of the cut-off. It may be that these contextual differences, rather than the consequences of the quotas, really explain any differences (or non-differences) in outcomes that I measure on either side of the cut-off. To make a convincing case that I am providing evidence about the effects of exposure to quotas per se, I need to account for the possible effects due to these other sources of variation.

My strategy for doing so is less than ideal, but I think it takes fullest advantage of what this admittedly imperfect natural experiment can tell us. I examine whether demobilization was associated with any adverse or positive shocks that might affect the ethnic sensitivities of subjects. Current theories of group threat and prejudice (Brewer, 1999) suggest that if there are adverse shocks associated with demobilization—e.g., a serious fall in material well-being or subjective perceptions of one's well-being—then the estimates based on the regression discontinuity alone would be biased in a manner that *overstates* the effect of participation in the integrated military on *reducing* prejudice and ethnic salience. The same theory suggests that positive shocks would bias the analysis in the other direction.

2.4.5 Outcomes and measurement

I study effects on two outcomes: prejudicial behavior and ethnic salience.

Prejudicial behavior

As discussed above, the debate over quotas suggests two contradictory hypotheses about the effects of participation in a quota-integrated institution on prejudicial behavior. The contact hypothesis suggests that the revision of prejudicial beliefs will cause a reduction in prejudicial tendencies. The resentment hypotheses suggests that such participation may increase prejudicial tendencies. This study evaluates which of these two effects tends to predominate among those participating in the quota-integrated military in comparison to those who were demobilized instead.

I measure impacts on prejudice as the extent to which participation in the integrated army *moderates* the effect of the enumerator's ethnicity on a respondent's willingness to respond to a set of politically or ethnically sensitive questions. Ethnic prejudice is a term that is used to mean many things, including the conditioning of behavioral responses conditional on ethnicity, or mere like or dislike of individuals of another ethnicity (Paluck and Green, 2009). The definition that I apply is the *conditioning of cooperation on ethnicity*. I assume that a survey respondent's decision to answer a survey question is an expression of cooperation. Moreover, the willingness to respond to sensitive questions is an expression of *trust* and *comfort* with the enumerator. Prejudice then refers to the conditioning of such trust and comfort, and thus cooperation, on the ethnic identity of the enumerator. I assume that the more prejudiced a person is, the more likely is that person to condition such trust and comfort, and thus cooperation, on the ethnicity of the enumerator, with higher trust and comfort with a co-ethnic enumerator than with a non-co-ethnic enumerator.[20]

I chose to focus on non-response to sensitive questions as a minimally intrusive

[20] The assumption is consistent with findings by Habyarimana et al (2007), who show that coethnics may be more likely to trust each other because they can expect that reciprocity will be upheld.

measure of prejudicial behavior. The "obtrusiveness" of a measure is based on the extent to which subjects under investigation are aware that their behavior is being assessed. Research in social psychology points clearly to the fact that subjects are reactive to knowledge that they are being assessed (Kazdin, 1979). Obtrusiveness compromises external validity, especially when measuring something like prejudice, because of social desirability considerations (Paluck and Green, 2009).

To construct the measure, I first calculate the degree of responsiveness to sensitive survey questions. This required determining a set of "sensitive" questions. A first cut list was developed based on discussions with our enumerators before and after fieldwork. The enumerators were not aware that this assessment of sensitivity would be used to construct a measure of prejudice. We did not change the questions either. Rather, we emphasized to the enumerators that at the beginning of each module in the questionnaire, they were to explain again that the respondent had the right to refuse to answer or indicate "I don't know" for any question posed.[21] I then narrowed the list to a set that exhibited substantial non-response. (Questions with little non-response provide no information, and so their exclusion does not contribute to bias.) The result was a list of 23 "sensitive" questions. (The questions are given in the appendix.) I then added three questions that were posed to the *enumerator* about whether they thought the respondent appeared "distracted," whether the enumerator was reluctant to answer questions, or whether the enumerator felt uncomfortable interviewing the respondent. I constructed a "non-responsiveness proportion" score by adding the number of questions to which a respondent failed to give an answer, plus the number of enumerator questions that were answered in the affirmative, and then dividing by the total, 26. Note that this measure, while rather easy to interpret, provides

[21]In constructing the non-response measure I do not distinguish between "refuse" and "don't know" responses, because as our enumerators explained, a common way for a respondent to dodge a question would be to say "don't know." This introduces a bit of measurement error. I see no reason to think that the error would bias the analysis, although it does contribute to some loss of statistical power.

equal weight to all questions. This may not be the optimal way to draw information on non-responsiveness, as there may be substantial correlation in the non-response likelihood for different clusters of questions. Thus, as a secondary measure, I used a two parameter logistic item response model to construct a factor score from the non-response patterns.[22]

The next step was to use the responsiveness scores to estimate *ethnic prejudice*. Our enumeration team consisted of seven enumerators who identified as Hutu, and 17 who identified as Tutsi. The ethnic distribution of our enumeration team differs greatly from the ethnic distribution of the population, in which Hutus are the large majority (popularly believed to make up 85% of the population).[23] Given that we recruited enumerators from among the university-educated, the disparity between our team's ethnic distribution and that of the population is a testament to the deep legacy of Hutu exclusion in the education system.[24] Our enumerator assignment protocol was such that enumerators were to be randomly assigned to respondents, who themselves were randomly selected. In this way, the protocol provided for the random assignment of enumerator-respondent pairs to be either ethnically *concordant* (both having the same ethnicity) or ethnically *discordant* (ethnicity of enumerator and respondent differ). Because we were not present to monitor enumerators at all times, there is some question as to whether enumerators may have violated this protocol, and sorted their interviews on the basis of ethnicity or some other factor that may confound the concordance/discordance assignment. I assessed this possibility by performing a randomization test. Our respondent data contain 1086 total interviews with still-active and demobilized FAB soldiers. Of them, 902 identify as Tutsi, and

[22]The model was fit with the "ltm" package in R, and factor scores were drawn using an empirical Bayes method (Rizopoulos, 2006). An ANOVA test indicated that a two-parameter model would be preferable to a one-parameter model.

[23]Refer to fn 5.

[24]Jackson (2000) provides an excellent look at ethnic and regional exclusion in Burundi's education system.

184 as Hutu. Given those proportions, as well as our 7 to 17 ratio of Hutu to Tutsi enumerators, purely random assignment would lead us to expect 64% of enumerator-respondent pairs to be ethnically concordant. In my sample, the figure is 65%. An exact binomial test indicates that probability of seeing this large of a departure given random assignment is 0.38, well within reasonable bounds. I think it is reasonable to believe that the scope for sorting on ethnicity was limited. We trained our enumerators extensively on the protocol, and emphasized that compensation and retention in the project would be based on their abiding by the protocol.

With the responsiveness scores and the concordance assignments set, prejudice is measured in terms of the difference in non-response rates across concordant and discordant pairs. In the total sample, the non-responsiveness score for ethnically concordant pairs is 0.12, while it is 0.13 for discordant pairs, suggesting only the slightest ethnic discrimination, if anything, in the overall sample. However, this small difference may mask substantial heterogeneity across groups, and it is this heterogeneity that I use to assess the impact of participation in the integrated military: the impact on prejudice is measured as the extent to which participation in military *moderates* prejudice, as measured via ethnic discordance effects. Effectively, I compute the ethnic discordance effect above and below the age eligibility cut-off. Then, I measure the impact of participation in the integrated army with a "difference in differences," estimating how the discordance effect differs for those above and below the cut-off. This difference is then weighted by the inverse complier proportion. In practice, this is done using a two-stage least squares estimator, where the interaction of ethnic concordance assignment and position relative to the eligibility cut-off is used to instrument the interaction of concordance assignment and participation in the military.

There are two technical issues that I also need to address in the analysis. First is that there may be non-linearities due to "ceiling" or "floor" effects, given that the raw non-response measure is truncated between 0 and 1. This would be especially problematic if the baseline non-response proportion (that is, the average non-response proportion among subjects who were not in an ethnically concordant pair) differed greatly on either side of the discontinuity. The logistic factor scores from the logistic item response model helps to alleviate this problem. The factor scores are a projection onto the logit scale, which in principle provides a linearizing transformation of a variable bounded between 0 and 1. Thus, the factor scores should provide for more reliable estimates of non-responsiveness effects in the case that baseline non-responsiveness rates differ significantly. Second, because the proportion of Hutu and Tutsi enumerators was not 50-50, the probability of being in a ethnically concordant or discordant pair varies depending on the ethnicity of the subject. For Tutsi subjects, the randomization procedure implies that the probability of being in a concordant pair is 17/24 and in a discordant pair is 7/24. For Hutu respondents, these probabilities are reversed. In order to ensure an unbiased estimate of the sample average effect of ethnic concordance, units are weighted by the inverse of their probability of being assigned to the concordance or discordance condition that they received.

Because the measure depends on the identity of the enumerator in the enumerator-respondent pair, it is useful to have more information on the profile of the enumerators. All enumerators were recruited through *Iteka–Ligue Burundaise des Droits de l'Homme* (Burundian League for Human Rights), a non-partisan, nationally prominent, and nationally syndicated human rights organization. The organization enjoys a reputation in Burundi of being a neutral advocate for human rights, taking stances at times in opposition of all political tendencies in the country. *Iteka* members hail from all parts of the country, although their operations are centered in the capital, Bujumbura. The enumeration staff included individuals from all regions of the country, although they were all recruited at *Iteka*'s central offices in Bujumbura. The enumerators ages ranged from 25 to 45, and included 8 women (of which 1 was Hutu). All enumerators had at least a university degree, which is a rare accomplishment in Burundi.[25]

[25]Samii (2010b) estimates that about 1% of Burundian civilians have a university degree.

Putting all of this together yields the following model for estimating effects of participation versus being demobilized on prejudicial behavior,

$$\begin{aligned}\text{Non-resposiveness}_i =& \beta_0 + \beta_1 \text{military}_i + \beta_2 \text{eth. conc.}_i + \beta_3 \text{military}_i * \text{eth. conc.}_i \\ &+ \beta_4(\text{age}_i - 45.5) + \beta_5 \text{below cutpoint} * (\text{age}_i - 45.5) \\ &+ \phi \text{higher order terms}_i + \epsilon_i,\end{aligned}$$

where "military$_i$" is an endogenous indicator variable for participation in the new military. I use "below cutpoint$_i$" as an excluded instrument for "military$_i$" and the interaction, "below cutpoint$_i$*eth. conc.$_i$" as an excluded instrument for "military$_i$*eth. conc.$_i$." The coefficient, β_3 is the estimate of the moderating effect on the ethnic concordance effect. Thus, β_3 measures the effect of participating in the military versus being demobilized on ethnically prejudicial behavior. The $(\text{age}_i - 45.5)$ term is centered at the cut-point, and so I fit the model using two-stage least squares, weighted to account for the unequal ethnic concordance probabilities. Inference is based on the usual frequentist least-squares principles, with standard errors asymptotically robust to clustering at the level of the sampling location.

I believe that this is a measure with a good deal of external validity, particularly relative to survey-question alternatives. Nonetheless, there are some reasons to believe that it may understate the true amount of prejudice. One reason is that, despite the myths, it is actually quite difficult to discern ethnicity based on appearance or name (which enumerators provided to respondents) in Burundi. For a given level of prejudice active in the mind of a respondents, erroneous judgments by the respondent as to the enumerator's ethnicity will cause the discordance effect to understate the actual amount of prejudice. I acknowledge this possibility. This possibility, combined with the rather small sample size and the fact that the measure requires the estimation of an interaction effect mean that this is a rather low powered test. But with no ability to increase the sample size, I accept this trade-off for the sake of improving the external validity of the measure.

Perceptions of ethnic salience

In addition to prejudice, ethnic conflict typically requires that ethnicity be understood as a salient feature of one's identity. Ethnic salience is related to "ethnocentrism," which involves not just co-identification, but also the tendency of individuals to view their own ethnic identity as requiring them to work in the interest of their co-ethnics in competition with other ethnic groups (Levine and Campbell, 1972; Brewer and Campbell, 1976). A combination of ethnocentrism and scarcity may result in ethnic identities defining salient lines of conflict (Brewer, 1999; Green and Seher, 2003; Posner, 2004). Such conflict may manifest itself overtly through practices of exclusion and outright inter-group fighting. There are also attitudinal manifestations, wherein masses of individuals perceive their ethnic identity as important, relative to other possible modes of self-identification, not only as a means of describing oneself, but also as a likely determinant of one's life prospects and an important marker on which to condition cooperation (Akerlof and Kranton, 2000; Chandra, 2006; Eifert et al, 2010; Fearon, 2003).

Brewer (1999:435-436) notes that a heightened sense of threat is likely to manifest itself is stronger expressions of ethnocentrism, and that this pattern has been documented in a study of ethnic attitudes among groups in South Africa. To the extent that soldiers perceive ethnic integration as threatening, we should record more intensely ethnocentric expressions. The contact hypothesis runs in precisely the opposite direction: it proposes that recategorization and the fostering of a common identity will result in our recording of less ethnocentric expressions. The soldiers studied in this paper provide an excellent test of these competing claims, as it is for FAB soldiers that quota-based integration represented a status demotion and, thus, a threat.

I use a set of three yes-or-no survey questions to measure perceptions of ethnic salience. The questions are as follows:

1. "According to me is it necessary to support ideas of other [respondent's ethnic

group] even if I do not fully agree with them."

2. "The wellbeing of [respondent's ethnic group] people in Burundi has more to do with politics than their own hard work."

3. "Things that happen to other [respondent's ethnic group] people in Burundi has an impact on my life."

"Yes" responses were coded as 1 and "no" responses as 0. I used the responses to construct a factor score, also based on a two-parameter logistic item response model.[26] Higher scores on the factor correspond to higher perceptions of ethnic salience. Responses to these scores exhibit missingness for 31 (3%), 57 (5%), and 62 (6%) respondents, respectively, with item missingness on at least one of the variables for 106 (10%). The missingness rate is low enough such that I simply omit the missing observations from the analysis. Note that these questions were not included in the construction of the non-responsiveness score. They were not considered particularly sensitive, and the low non-response rates attest to that.

Thus, the model that I estimate in this case is given by,

$$\begin{aligned}\text{Ethnic salience}_i =& \alpha_0 + \alpha_1 \text{military}_i + \alpha_2(\text{age}_i - 45.5) + \alpha_3 \text{below cutpoint} * (\text{age}_i - 45.5) \\ &+ \phi \text{higher order terms}_i + \epsilon_i,\end{aligned}$$

where again "military$_i$" is an endogenous indicator variable for participation in the new military, "below cutpoint$_i$" is used as an excluded instrument for "military$_i$," and $(\text{age}_i - 45.5)$ is centered at the cut-point. I fit the model using two-stage least squares. In this case, the coefficient, α_1, measures the effect of participation in the military on ethnic salience. Again, inference is based on the usual frequentist least-squares principles, with standard errors asymptotically robust to clustering at the level of the sampling location.

[26] See fn. 22.

2.4.6 Examining the cross-validity of the outcome measures

[Figure 2.3 here.]

Figure 2.3 shows the relationship between the non-response-based prejudical behavior measure and the salience score for subjects with ages within 10-year and 5-year windows around the cut-point. The upper (red) line in both graphs is the regression fit for subjects in the ethnically discordant interview condition, and the lower (blue) line is the regression fit for subjects in the ethnically concordant condition. A few things are apparent. First, we do see that predicted marginal relationships are evident: the non-response proportions rise (albeit slightly) in levels of salience, and we see that those in the ethnically discordant condition have higher rates of non-response. The "salience" effect is not significant at even the 60% level, although the concordance effect is significant at the 6% level in the 10-year window, and only at the 23% level in the 5-year window. Second, were it that the prejudicial behavior and salience measures were measuring a common underlying trait, we would expect that the magnitude of the discordance/concordance effect would be larger at higher levels of salience. This is not evident, as there is no apparent interaction effect (from the regressions, the interactions are not significant at even the 90% significance). Thus, the ethnic concordance/discordance effects and salience scores are both picking up on latent sources of reluctance to answer sensitive questions, although in ways that are partially distinct. This is not as tidy a result as one might want, although it may be indicative of the complexity of the relationship between latent attitudes, behavior, and expressed attitudes when it comes to prejudice and ethnic salience.

2.5 Results

2.5.1 Effects on prejudicial behavior

[Figure 2.4 here.]

Figure 2.4 graphs the raw data (jittered to show masses of data) for the non-response proportions and non-response factor scores in the left column, and then plots estimates of the effect of ethnic concordance in enumerator-respondent pairs over the age variable in the right column. The gray areas in the graphs show the region below the cut-point, demarcated by 0 on the lower axis. These outcomes are "unadjusted" in that they do not account for the fact that treatment status (whether one is a member of the military or not) is not determined perfectly by the cut-point. The two-stage least squares regressions that I discuss below perform that adjustment. The ethnic concordance effects graphed in the right column of Figure 2.4 were estimated by with ordinary least squares regressions within each of the 5-year bins demarcated in the graphs. The dots show the point estimates of the concordance effect, and the bars show 90% confidence intervals from the OLS regressions. The quantity of interest here is the amount that the ethnic concordance effect varies on either side of the cut-point. We observe a substantial downward jump in the estimated effect. For those just to the left of the cut-point, the ethnic concordance of the enumerator-respondent pair has little or no effect, whereas for those just to the right of the cut-point, concordance has a substantial effect. In other words, prejudicial behavior is greater among the latter.

As Imbens and Lemieux (2007) suggest, it is important to examine whether there are jumps elsewhere in the graph, away from the cut-point. While such jumps do not invalidate the causal identification strategy per se, significant jumps elsewhere in the graph may cast doubt on whether what we see near the cut-point is genuine evidence of a causal effect. We note that there is a slight deviation in effect size trends for those in the bin between -10 and -5 years from the cut-point. This deviation is not nearly as large as the jump at the cut-point, but it does beg for some explanation. One possibility comes if we return for a moment to Figure 2.2. We note that going from the $[-10, -5)$ bin to the $[-5, 0)$ bin, there is a substantial drop in the proportion of military versus demobilized, although the drop does not deviate terribly from the overall trend. Nonetheless, to the extent that prejudicial behavior is the result of some interactive relationship between age and demobilization status, we might expect that the graph of prejudicial behaviors over age to exhibit a curvilinear relationship, given that differences between members of the military and demobilized soldiers are being integrated in different proportions. The only other major deviation that we see is in the highest bin, $[10, \infty]$, but the noisiness of the estimate for this bin means that little can be inferred.

To quantify the effects of participation in the integrated military on prejudicial behavior, I use two-stage least squares (TSLS) regression. The results are displayed in Table 2.1.[27] Because the age variable is centered at the cut-point, the coefficient "Military" estimates the local (i.e. at the cut-point) average treatment effect of participation in the military for respondents in ethnically discordant pairs, the coefficient on "Ethnic concordance" estimates the local average treatment effect of ethnic concordance for those not in the military, and the coefficient on "Military $*$ Ethnic concordance" estimates the local average moderating effect of being in the military on the effect of ethnic concordance. It is the latter that we are interested to test the hypotheses about prejudicial behavior. I present estimates for the raw non-responsiveness proportions in the first three columns of estimates, and then for the non-responsiveness factor score in the last three columns. As discussed above, the former are easier to interpret, but the latter make more efficient use of the information. Estimates are given for local linear regressions within a 3-year window and 5-year window, and a quadratic regression within the 10-year window. For the latter, the specification search strategy outlined above was used, and no higher order coefficients had p-values less than 0.10.

The contact hypothesis would lead us to expect the that effects of ethnic concordance would be weaker among the members of the integrated military, implying

[27]Note that the strong first stage with respect to treatment assignment has already been demonstrated in the section above on "Treatment definition."

less prejudicial behavior in the military. The resentment hypothesis would lead us to expect that the concordance effect would be stronger. The 5-year and 10-year estimates show that the concordance effect is significantly less than zero for those not in the military, indicating measurable prejudicial behavior. However, the coefficient on "Military $*$ Ethnic concordance" effectively cancels out this effect, implying that among members of the military, ethnic concordance has no effect. For the 5-year and 10-year windows, the p-values on the interaction effect are quite small, suggesting strong evidence. The 3-year window estimates are extremely noisy, but the point predictions are quite consistent with the 5-year and 10-year window estimates. Estimates using a triangular kernel also result in very similar point estimates, although as expected, the p-values are substantially larger. Thus, I find that the data are consistent with the expectations of the contact hypothesis: participation in the integrated military is associated with a reduction in the general prevalence of prejudicial behavior in comparison to those who do not participate.

The estimates in Table 2.1 show that the estimated level of nonresponsiveness among military participants at the cut-off is substantially lower among those in ethnically discordant enumerator-respondent pairs than is the case for their demobilized counterparts. This may raise concern about other lurking differences on either side of the cut-off that may also be confounding the analysis. In order to investigate this, I examined plots of the non-response proportions and non-responsiveness factor scores within the 3-year and 5-year bandwidths, and placed over these plots linear regression fits within the 3-year and 5-year bandwidths. (These are shown in Figure 2.7 in the Appendix.) What one discovers is that apparent differences in "baseline" levels of non-responsiveness are in part due to the fact that the ethnically *discordant* condition is being coded as the "baseline" condition. However, if we reverse this coding and set the ethnically *concordant* condition as the baseline condition, then the two baselines do line up quite well. (This is evident in looking at the position of the red regression lines on either side of the cut-point in Figure 2.7 in the Appendix.) Thus, I re-ran the regressions that appear in Table 2.1, but I used "1−eth. conc." as the independent variable, and adjusted the inverse probability weights appropriately. The results (Table 2.5 in the Appendix) were such that the coefficient on "military" is considerably smaller than was the case in Table 2.1, and in no cases is the coefficient significant at conventional levels. The coefficients on the "military X ethn. conc." interaction term does not change appreciably, and with the exception of the first column estimate, obtain p-values below 0.10. Thus, the re-coding has us maintain our conclusion that there is an appreciable moderating effect on the non-response proportion while also showing that these differences are not a result of significantly different "baseline" responses.

2.5.2 Effects on ethnic salience

[Figure 2.5 here.]

Figure 2.5 graphs,in the left panel, the raw data for the ethnic salience factor scores by age, and then in the right panel average values within the five-year bins marked on the graph. Visual inspection of the graphs shows no evidence of a substantial jump at the cut-point. Rather, what one observes is a rather smoothly increasing relationship, starting at the $[-5, 0)$ bin and continuing through the cut-point into higher age values. We do notice that the slope relating age and ethnic salience is flat for the age range below the cut-point but begins to rise steeply above the cut-point. This may be indicative of a "suppression effect" among those in the military, although the causal identification strategy that I am use here does necessary permit such inference.

Table 2.2 presents TSLS estimates for the effect of participation in the military on ethnic salience scores, adjusting for the probabilistic nature of treatment assignment on either side of the cut-point. The coefficient on "Military" is our estimate of the local average treatment effect on the ethnic salience score. There is nothing to suggest a substantial or significant effect either way. We do note that the estimates on the

"Below cutpoint ∗ Age" interaction are suggestive of the dampening effect that we observed from the graphs, at least for the 3-year and 10-year windows. But again, the identification strategy does not provide us with anuy special reason to attribute this to participation in the military per se. It may very well be that had those participants below the cut-point not participated in the military, that their ethnic salience scores would have looked the same. Results using a triangular kernel are very similar, and no conclusions change. While the components of the ethnic salience score each exhibit low rates of missingness (less than 6% for all three component variables), in combination they result in a missingness rate of just under 10% for the salience score. I thus estimated the same models as above, but controlled for "ethnic concordance," as we know that this is related to missingness. The coefficients change slightly, but the standard errors grow as well, and in no case is the p-value on the military term less than 0.5. Thus, the evidence with respect to ethnic salience provides no indication that effects anticipated by either the contact or resentment hypotheses tend to dominate.

2.5.3 Enumerator fixed effects

Chance imbalances in the distribution of particular enumerators or enumerator traits may confound the analysis if enumerator-specific effects are present. In order to assess this I re-estimated all of the models above with fixed effects included for each estimator. These estimates will be quite noisy, as fixed effects estimation here requires that we partial out mean values for each of 24 enumerators in rather small samples. (Results are shown in Tables 2.6 and 2.7 in the Appendix.) Three of the surveys in the dataset did not have enumerator identification information entered, and so these were simply omitted. For the regressions of the prejudicial behavior measures, estimates become appreciably noisier and the coefficient estimates on the "eth.conc." term and the "military X ethn. conc." interaction term are slightly closer to zero.[28] Thus,

[28] F-tests reject the hypothesis of jointly insignificant fixed effects in all models.

controlling for enumerator fixed effects weakens the statistical conclusions drawn from the unadjusted analysis for the analysis of prejudicial behavior, although it does not cause us to reverse our conclusion completely. The coefficients on the salience scores change rather erratically for the regression in the "optimal" 4-year bandwidth and change only slightly for the other two. In none of the cases would we reject the null at reasonable significance levels, and so there is no change in our inferences.

2.6 "Placebo" tests using pre-treatment variables

In addition to the checks that we performed above on the behavior of the treatment variable and the outcome variables, Imbens and Lemieux (2007) recommend conducting "placebo" tests by analyzing variables that could not possible have been causally affected by treatment. One wants to do this on pre-treatment variables that have strong potential to confound were they to exhibit discontinuities near the cut-off. I located five such variables in the survey data:

Non-commissioned officer status Our data consist of rank-and-file and NCOs. We might imagine that NCOs, being of higher rank, would be more likely to react with resentment to the integration of "irregular" rebel forces into the new military.

Years in the military A similar argument as above may be said for those with more time in the military.

Years of education, pre-war Years of education may be associated with a more tolerant world view, on the basis that ignorance and intolerance go together. Alternatively, pre-war education levels are a reliable measure of socio-economic status prior to integration. It may be that higher socio-economic status is associated with less tolerance in post-war Burundi, as it is those who were more privileged before the war that face a greater "threat" from the redistributive changes brought about as a result of the war.

Unit death rate Those from units that suffered higher death rates may be more ethnically intolerant, as violence during the war was ethnically colored.

Family death rate The same argument as above may apply for those whose families suffer higher rates of wartime mortality.

Table 2.3 shows the results of these placebo tests on pre-treatment variables. Tests are shown for the preferred window (5-year) from the analysis above, based on TSLS regressions using the same specification as used in the analysis of ethnic salience scores. For none of the variables do we find compelling evidence of a confounding discontinuity. Models fit to either 3-year or 10-year bandwidths (not shown), with the latter using higher order terms, do not yield different results.

2.7 Alternative interpretations

[Figure 2.6 here.]

The results above show that when we compare those who participated in the new military versus being demobilized, we see less ethnically prejudicial behavior on average and no apparent change in expressions of ethnic salience. Relating this to our discussion above, one may conclude that there is more to recommend the optimistic "contact" theory about the effects of quota-based ethnic integration over the more pessimistic theories of hierarchy maintenance or "freezing" of conflictual identities. These results are in line with finding from hundreds of other studies, as reviewed in Pettigrew and Tropp (2006).

However, as discussed above, this natural experiment is not "clean" in its identification of the effects of quota-based integration *per se*. This is depicted in Figure 2.6. The figure displays graphs of two causal pathways. Graph (i) on the left displays a causal pathway showing that participation versus demobilization may cause changes in other relevant variables that ultimately affect expressions of ethnic prejudice or salience. It may be that it is these other relevant changes that are important, and that the pathway that flows through "exposure to integration" is of little importance. For example, participation in the military may cause exposure to military norms or training that would mute expressions of prejudice or salience even were there no exposure to integration. Alas, the current data do not allow me to assess this possibility. Alternatively, demobilization may result in changes to one's material well-being or psychological state that heighten expressions of prejudice or salience, in which case our "optimistic" interpretation of the findings above may be invalid. This is something that I can study to a certain extent, and I do so below. Graph (ii) on the right displays an alternative causal pathway. Here, exposure to integration is the primary thing that determines whether military participants will differ from their demobilized counterparts in their expressions of ethnic prejudice or salience. It may also be the case that such exposure affects other outcomes, but this is of no consequence for our interpretation. If the latter graph accurately characterizes what is actually happening, then the "optimistic" interpretation of the findings thus far is valid. The latter graph entails an "exclusion restriction," whereby there are no pathways circumventing "exposure to integration" that link participation/demobilization to expressions of ethnic prejudice or salience. In reality, it is unreasonable to believe that the exclusion restriction holds exactly. While the discontinuity ensures that exposure to the quota was not due to *self-selection*—a hugely important step in narrowing the scope for confounding—there are certainly other differences in military versus demobilized life that have nothing to do with exposure to ethnic integration. The question is whether these other differences are of sufficient magnitude to overturn the "optimistic" interpretation of the findings thus far.

The problem is that the data do not allow us to measure exclusion restriction violations. The situation is very similar to the classic "mediation" problem in statistics, and as Green et al (2010) have demonstrated, isolating exogenous variation in a treatment variable (in our case, participation versus demobilization) is not sufficient to identify whether treatment effects operate via one or another causal pathway. The robust solution to the problem is nothing short of a new study based on exogenous variation on precisely the factor of interest. In principle, this could be done for the current analysis, if for example ethnic integration was randomly phased-in, unit by unit, and one had data on members of both integrated and non-integrated units. But such is not the case with the data available.[29] Thus, I am limited to second-best tests that merely allow me to assess the *plausibility* of whether any exclusion violations are sufficiently mild as to retain the current interpretation. The tests involve studying whether there are substantial differences in other, potentially confounding post-treatment outcomes around the cut-off. If the evidence shows that there are *no* such differences, then the exclusion violation shown in graph (i) is less plausible, in that it greatly restricts the range of possibilities for how such an exclusion violation may occur.[30]

I test whether there are any changes in subjective or material well-being at the cut-off. One may propose that those who have been demobilized may now struggle more for subsistence than those who remain in the military. As Brewer (1999) notes, such vulnerability may translate into a stronger need for in-group solidarity. To the extent that one views this uncertainty in terms of competition between groups

[29] This is something I am trying to look into for future research in other post-conflict transition processes.

[30] As Green et al (2010) show, an estimated average "non-relationship" cannot distinguish between exclusion and the possibility that heterogenous causal effects still provide for an operative causal pathway that violates exclusion. For example, suppose that demobilization causes higher income for some and lower income for others. But, suppose that for those whose income decreases from demobilization, this negative income shock increases prejudice. In contrast, for those whose income increases from demobilization, this *positive* shock increases prejudice. Then, demobilization's average effect on income may be about zero, but this causal pathway is operative in linking demobilization to prejudice.

over scarce resources, this may in turn translate into heightened prejudice toward out-groups. To assess subjective perceptions, I used responses from the survey to a question of whether respondents considered their current economic conditions to be very bad, bad, good, or very good. Given that there were very few responses in either the very bad or very good categories (about 6% of responses in total), I constructed a binary variable that distinguished very bad and bad responses from very good and good responses. I applied the same specification as used above for the analysis of salience and the placebo checks, except that I also included the "ethnic concordance" treatment variable in the specification, as there was a non-trivial amount of missing data (see the summary statistics table in the Appendix). The estimates show no substantial differences at the cut-point in these perceptions (Table 2.8 in the Appendix). I should say, however, that the survey data do contain demobilized soldiers' responses to a question asking about whether they think things are better or worse for them as compared to their counterparts who remain in the military. Looking only at the responses of demobilized soldiers within the window of 45-50 years of age (within 5 years of the cut-point), 56 out of the 61 demobilized soldiers in this subgroup indicated "worse" (data were missing for 3 of the respondents), suggesting some bitterness among this group. This question was only asked of demobilized soldiers, however, so I cannot use it to construct a test for exclusion restriction violations. That said, members of the military and demobilized soldiers alike respond frequently that their current economic conditions are "bad" or "very bad." Limiting the analysis again to respondents within 5-years of the cut-point, 50% of military respondents indicated that their economic conditions were "bad" or "very bad," while the percentage was 58% for demobilized.[31] One could chalk up demobilized soldiers' apparent bitterness about not being the military to general dissatisfaction with their economic conditions—a dissatisfaction that is also strongly felt among current soldiers.

[31] A Chi-square test fails to reject the null at 33% significance.

To assess possible exclusion violations due to changes in material conditions, I estimated changes in the natural log of (self-reported) monthly income at the cut-point, again using the specification from the analysis of salience and placebo outcomes.[32] Income is an interesting outcome to examine, in this case, because there is no reason to believe that it would be affected by exposure to ethnic integration within the military. Thus, any differences that we measure must be due to income shocks associated with being demobilized versus remaining in the military. The estimates show no such shock at the cut-point (Table 2.9 in the Appendix). This is to be expected for reasons specific to the Burundi case. Demobilized soldiers from the national army in Burundi were afforded a combination of pension benefits and "reintegration" assistance. An income allowance was provided so as maintain a subsistence level comparable to that of military members for 2 years after being discharged. In addition, a World Bank supported reintegration program provided demobilized soldiers with financial capital, start-up materials, and training for establishing a civilian livelihood.[33]

2.8 Conclusion

A crucial issue in political development is why ethnic conflict endures and what actions might be taken to transcend it. It is incumbent upon political scientists to study policies that might be used to achieve such transcendance. Among the

[32]The rate of missingness was about 5% overall for the income measure, and so I simply omit those observations

[33]Gilligan et al (2010) study the impact of this assistance program on the economic and political reintegration of former rebel soldiers. Former national army members qualified for this same assistance as well as additional perks due from the national army pension scheme. Mvukiyehe et al (2006), Uvin (2007), and Verwimp and Bundervoet (2008) have demonstrated, from different data sources, that the reintegration benefits were quite generous, with demobilized soldiers enjoying a substantially higher standard of living than their civilian counterparts and at a level comparable to their military counterparts.

policy options, ethnic quotas are common features of transitional agreements. Despite this, little empirical work has been done to measure the effects of such policies on levels of ethnic conflict, although there exists much theorizing about possible effects. Remarkably, the theories point in two opposite directions, with one group of theories optimistic about the how quotas may reduce conflict, and another pessimistic about how they may exacerbate conflict.

In this study, I take advantage of an interesting natural experiment that arose in the context of the military integration process in post-war Burundi. The peace agreement that brought about an end to over a decade of terrible, ethnically charged violence called for the ethnic integration of the military according to a 50-50 quota. In practice, this was achieved largely by integrating members of the mostly Hutu rebel forces with members of the mostly Tutsi national army. The integration process saw the full mixing of former rebel and national army members into mixed units who trained, lived, and fought together.

As part of the process, an age cut-off was used to determine eligibility. I use this cut-off as a source of causal identification for the effects of participation in the integrated army versus being demobilized. I examine effects on prejudicial behavior and perceptions of ethnic salience among members of the defunct national army. I test a hypothesis derived from the literature on "inter-group contact," which suggests that enhanced contact due to integration should reduce prejudice and ethnic salience, against theories of "resentment" and "freezing" of conflictual identities, which suggest that integration may increase prejudice and salience. To measure prejudicial behavior, I use an unobtrusive measure based on rates of non-response to ethnically and politically sensitive questions conditional on the ethnicity of the enumerator relative to the respondent. To measure salience, I use questions that get at subjects' perceptions about whether their ethnic identity is important for conditioning their behavior and interpreting their life prospects. Using a "fuzzy" regression discontinuity design, I estimate that participation in the integrated army is associated with less prejudicial behavior, consistent with the contact hypothesis. This finding is also consistent with a large body of research on the effects of "optimal" contact between groups (Pettigrew and Tropp, 2006). I find no substantial effects on ethnic salience, suggesting that the changes in prejudicial behavior may not have to do with the reconceptualization of one's own identity. To make the case that these estimates capture the effects of integration *per se*, I examine whether there is evidence of confounding due to other possibly important variables near the cut-off. I do not find evidence of discontinuities in potentially confounding pre-treatment variables at the cut-point. Neither do I find evidence that for confounding due to shocks to well-being associated with being demobilized. I cannot rule out that other features of the experience of being in the new military may be at work independently of inter-ethnic contact. Nonetheless, my judgment that this is a large factor in explaining what we see in the evidence is based on my appreciation that integration and resulting forced inter-ethnic contact are a dominant factor that distinguishes the experience of former national army members who participated in the new army as opposed to those who did not.

I hope this study will stimulate greater interest in studying the micro-foundations of quota-based integration as a strategy to address legacies of exclusion and ethnic conflict. This study only looks at one small piece of a much larger puzzle. Do these effects endure? To what extent do the local effects that I measure here reverberate societally? These are questions that future research ought to address. I have also taken great lengths to explain how the natural experiment in this paper is far from perfect, but that obvious possible confounders do not seem to taint the analysis. It should be clear what kinds of situations would allow for a more refined analysis. Institutional integration processes often unfold in stages. If such a process can be designed such that these units within an institution are subject to integration in stages in a more or less random manner, then one may be able to assess the effects of such integration more cleanly. Researchers should seize on such opportunities. Also, I should clarify that the research design here does not allow me to study whether quota-based integration is *more or less* effective than other strategies—e.g., strategies that purposefully avoid using ethnic categories—in reducing ethnic conflict. That would require comparison to cases where these other types of strategies are employed. All I can do here is to comment on whether presumed micro-effects of reduction or exacerbation of ethnic conflict are manifest in response to exposure to involuntary ethnic integration. In addition, this paper has attempted to develop an unobtrusive measure of prejudice. The measure here was improvised during the surveying process, and the approach could certainly be improved. While I think the measure is credible, clearly more thought could go into designing questions to produce more reliable variation in non-response patterns and also to better make use of enumerator identity to stimulate latent tendencies in respondents. There would seem to be ample room for experimentation, perhaps drawing on results from recent experiments on "race-of-interviewer" effects.[34]

As for policy implications, quota-based integration of institutions is a commonly employed strategy in addressing legacies of ethnic exclusion and ethnic conflict. The appeal of quotas is based on their transparency as well as on the appreciation that legacies of exclusion may not disappear without quotas or affirmative action (Fryer and Loury, 2005; Duflo, 2005). Of course, there are principled objections to quotas. But another important source of objection is the claim that they may have pernicious effects that exacerbate the very conflict that they are ostensibly meant to transcend. To the extent that the evidence in this paper characterizes more general phenomena, these findings weigh against that latter objection.

[34]For recent contributions, with references to classic studies, see Davis and Silver (2003) and Krysan and Couper (2003).

References

Aitken R. "Cementing divisions? An assessment of the impact of international interventions and peace-building policies on ethnic identities and divisions." *Policy Studies*. 28(3):247-267.

Akerlof GA, Kranton RE. 2000. "Economics and identity." *The Quarterly Journal of Economics*. 115(3):715-753.

Allport GW. 1954. *The Nature of Prejudice*. Reading: Addison Wesley.

Blumer H. 1958. "Race prejudice as a sense of group position." *Pacific Sociological Review*. 1:3-7.

Brewer MB. 1999. "The psychology of prejudice: Ingroup love or outgroup hate?" *Journal of Social Issues*. 55(3):429-444.

Brewer MB, Campbell DT. 1976. *Ethnocentrism and Intergroup Attitudes: East African Evidence*. Beverly Hills: Sage.

Chandra K. 2006. "What is ethnic identity and does it matter?" *Annual Review of Political Science*. 9:397-424.

Chauchard S. 2010. "Can the experience of political power by a member of a stigmatized group change the nature of day-to-day interpersonal relations? Evidence from rural India." Typescript, New York University.

Coser L. 1956. *The Functions of Social Conflict*. New York: Free Press.

Davis DW, Silver BD. 2003. "Stereotype threat and race of interviewer effects in a survey of political knowledge." *American Journal of Political Science*. 47(1):33-45.

Duflo E. 2005. "Why political reservations?" *Journal of the European Economic Association*. 3(2):668-678.

Dunning T. 2010. "Do quotas promote ethnic solidarity? Field and natural experimental evidence from India." Typescript, Yale University.

Eggers AC, Hainmueller J. 2009. "MPs for sale? Returns to office in postwar British Politics." *American Political Science Review*. 103(4):1-21.

Eifert B, Miguel E, Posner DN. 2010. "Political competition and ethnic identification in Africa." *American Journal of Political Science*. 54(2):494-510.

Fearon JD. 2003. "Ethnic and cultural diversity by country." *Journal of Economic Growth*.

Fryer RG, Loury GC. 2005. "Affirmative action and its mythologies." *Journal of Economic Perspectives*. 19(3):147-162.

Gibson JL, Gouws A. 2003. *Overcoming Intolerance in South Africa*. Cambridge: Cambridge University Press.

Gilligan M, Mvukiyehe E, Samii C. 2010. "Reintegrating rebels into civilian life: Quasi-experimental evidence from Burundi." Typescript, Columbia University and New York University.

Green DP, Seher RL. 2003. "What role does prejudice play in ethnic conflict?" *Annual Review of Political Science*. 6:509-531.

Green DP, et al. 2009. "Testing the accuracy of regression discontinuity analysis using experimental benchmarks." *Political Analysis*. 17(4):400-417.

Green DP, Ha SE, Bullock JG. 2010. "Enough already about "black box" experiments: Studying mediation is more difficult than most scholars suppose." *The Annals of the American Academy of Political and Social Science*. 628(1):200-208.

Gurr TR. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.

Habyarimana J, Humphreys M, Posner DN, Weinstein JM. 2007. "Why does ethnic diversity undermine public goods provision?" *American Political Science Review*. 101(4):709-725.

Hahn J, Todd P, Van Der Klaauw W. 2001. "Identification and estimation of treatment effects with a regression-discontinuity design." *Econometrica*. 69(1):201-209.

Horowitz DL. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Horowitz DL. 2001. *The Deadly Ethnic Riot*. Berkeley: University of California Press.

Hutchful E. 1997. *The Military and Militarism in Africa*. Dakar: CODESRIA.

Imbens GW, Kalyanaraman K. 2009. "Optimal bandwidth choice for the regression discontinuity estimator." NBER Working Paper 14726.

Imbens GW, Lemieux T. 2007. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics*.142(2):615-635

Jackson T. 2000. *Equal Access to Education: A Peace Imperative in Burundi*. London: International Alert.

Kazdin AE. 1979. "Unobtrusive measures in behavioral assessment." *Journal of Applied Behavior Analysis*. 12:713-724.

Krebs R. 2004. "A school for the nation? How military service does not build nations, and how it might." *International Security*. 28(4):85-124.

Krysan M, Couper MP. 2003. "Race in the live and the virtual interview: Racial deference, social desirability, and activation effects in attitude surveys." *Social Psychology Quarterly*. 66(4):364383.

Lee DS. 2008. "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics.* 142(2):675-97.

Lemarchand R. 1970. *Rwanda and Burundi.* London: Pall Mall.

Levine RA, Campbell DT. 1972. *Ethnocentrism: Theories of Conflict, Ethnic Attitudes, and Group Behavior.* New York: John Wiley.

Mvukiyehe E, Samii C, Taylor G. 2006. *Stabilizing the postwar environment in Burundi: Preliminary results and recommendations from a 2006 pilot survey.* Typescript, Columbia University.

Paluck EL, Green DP. 2009. "Deference, dissent, and dispute resolution: An experimental intervention using mass media to change norms and behavior in Rwanda." *American Political Science Review.* 103:622-644.

Paluck EL, Green DP. 2009. "Prejudice reduction: What works? A critical look at evidence from the field and the laboratory." *Annual Review of Psychology.* 60:339-367.

Petersen R. 2002. *Understanding Ethnic Violence.* Cambridge: Cambridge University Press.

Pettigrew TF. 1998. "Intergroup contact theory." *Annual Review of Psychology.* 49:65-85.

Pettigrew TF, Tropp LR. 2006. "A meta-analytic test of intergroup contact theory." *Journal of Personality and Social Psychology.* 90(5):751-783.

Posner DN. 2004. "The political salience of cultural difference: Why Chewas and Tumbukas are allies in Zambia and adversaries in Malawi." *American Political Science Review.* 98(4):529-545.

Rizopoulos D. 2006. "ltm: An R package for latent variable modeling and item response theory analyses." *Journal of Statistical Software*. 17(5).

Samii C. 2010a. "Military integration in Burundi." In Licklider R. Eds. *New Armies from Old: Merging Competing Military Forces after Civil War*. Forthcoming.

Samii C. 2010b. "Who Wants to Forgive and Forget? Transitional Justice Preferences in Post-war Burundi."' Paper presented at APSA, September 2010.

Simonsen SG. 2005. "Addressing ethnic divisions in post-conflict institution-building: Lessons from recent cases." *Security Dialogue*. 36(3): 297-318.

Spilerman S. 1970. "The Causes of Racial Disturbances: A Comparison of Alternative Explanations." *American Sociological Review* 35:627-49.

United Nations. 1996. *Final Report of the International Committee of Inquiry for Burundi*. New York: United Nations.

Uvin P. 2007. *Ex-combatants in Burundi: Why they joined, why they left, how they fared*. World Bank MDRP Working Paper, No. 3.

Verwimp P, Bundervoet T. 2008. *Consumption growth, household splits, and civil war.* Households in Conflict Working Paper, No. 48.

Wimmer A. 2006. "Ethnic exclusion in nationalizing states." In Delanty G, Kumar K, eds. *Handbook of Nations and Nationalism.* London: Sage.

Wolpe H, McDonald S. 2006. "Burundi's transition: Training leaders for peace." *Journal of Democracy.* 17(1):126-132.

World Bank. 2004. *Technical Annex for a Proposed Grant of SDR 22.2 Million to the Republic of Burundi for an Emergency Demobilization, Reinsertion and Reintegration Program.* Washington, DC: World Bank.

Tables

Table 2.1: **TSLS regressions of non-response proportions and non-response factor scores**

	(1)	(2)	(3)	(4)	(5)	(6)
	Non-response proportion			Non-response factor score		
Military	-0.25	-0.11**	-0.12**	-1.02**	-0.81***	-0.28
	(0.15)	(0.05)	(0.06)	(0.43)	(0.31)	(0.25)
MilitaryXeth. conc.	0.07	0.09*	0.06*	0.66**	0.64**	0.53***
	(0.09)	(0.05)	(0.04)	(0.30)	(0.28)	(0.18)
Eth. conc.	-0.06	-0.08**	-0.07**	-0.46**	-0.49***	-0.57***
	(0.05)	(0.03)	(0.03)	(0.19)	(0.17)	(0.14)
Age (centered)	-0.05	-0.01	-0.03	-0.15	-0.06	-0.01
	(0.05)	(0.02)	(0.02)	(0.10)	(0.07)	(0.03)
Below cutpointXAge (centered)	-0.01	0.01	0.02	0.05	-0.03	0.02
	(0.03)	(0.02)	(0.03)	(0.09)	(0.09)	(0.04)
Age (centered) sq			0.00			
			(0.00)			
Below cutpointXAge (centered) sq			-0.00*			
			(0.00)			
Constant	0.29**	0.22***	0.24***	0.62*	0.47*	0.33*
	(0.11)	(0.05)	(0.06)	(0.33)	(0.27)	(0.20)
N	110	160	264	140	160	264

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (1) uses an Imbens-Kalyanamaran optimal bandwidth of 3 years, and models (2) and (3) use 5-year and 10-year bandwidths, respectively. Model (4) uses an Imbens-Kalyanamaran optimal bandwidth of 4 years, and models (5) and (6) use 5-year and 10-year bandwidths, respectively.

Table 2.2: **TSLS regressions of ethnic salience factor scores**

	(1)	(2)	(3)
	Salience factor score		
Military	0.00	-0.04	-0.14
	(0.41)	(0.31)	(0.16)
Age (centered)	0.03	0.03	0.04**
	(0.09)	(0.07)	(0.02)
Below cutpointXAge (centered)	0.04	0.01	-0.04*
	(0.10)	(0.08)	(0.02)
Constant	0.05	0.04	0.05
	(0.25)	(0.20)	(0.10)
N	125	142	238

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (1) uses an Imbens-Kalyanamaran optimal bandwidth of 4 years, and models (2) and (3) use 5-year and 10-year bandwidths, respectively.

Table 2.3: **TSLS regressions of pre-treatment ("placebo") variables**

	(1)	(2)	(3)	(4)	(5)
	NCO	Military yrs.	Pre-war ed.	Unit death rt.	Family death rt.
Military	0.15	4.48	0.92	-0.06	-0.03
	(0.13)	(3.57)	(0.98)	(0.07)	(0.10)
Age (centered)	0.03	1.47*	0.32	-0.00	0.00
	(0.03)	(0.89)	(0.25)	(0.01)	(0.03)
Below cutpointXAge	-0.03	-0.40	-0.25	-0.02	0.00
	(0.02)	(0.91)	(0.27)	(0.02)	(0.03)
Constant	0.88***	21.75***	6.66***	0.09*	0.17**
	(0.10)	(2.76)	(0.67)	(0.05)	(0.07)
N	138	152	160	143	160

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All models are fit using the 5-year bandwidth.

Figures

Figure 2.1: **Sample proportions of military members of different ethnicities in sampled barracks**

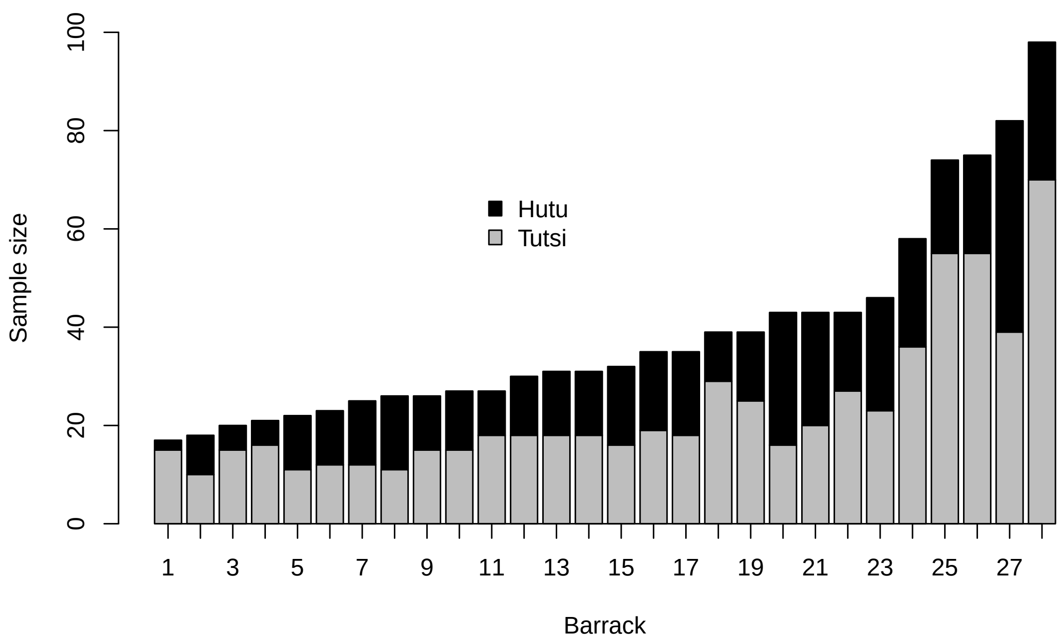


Figure 2.2: **Treatment assignment**

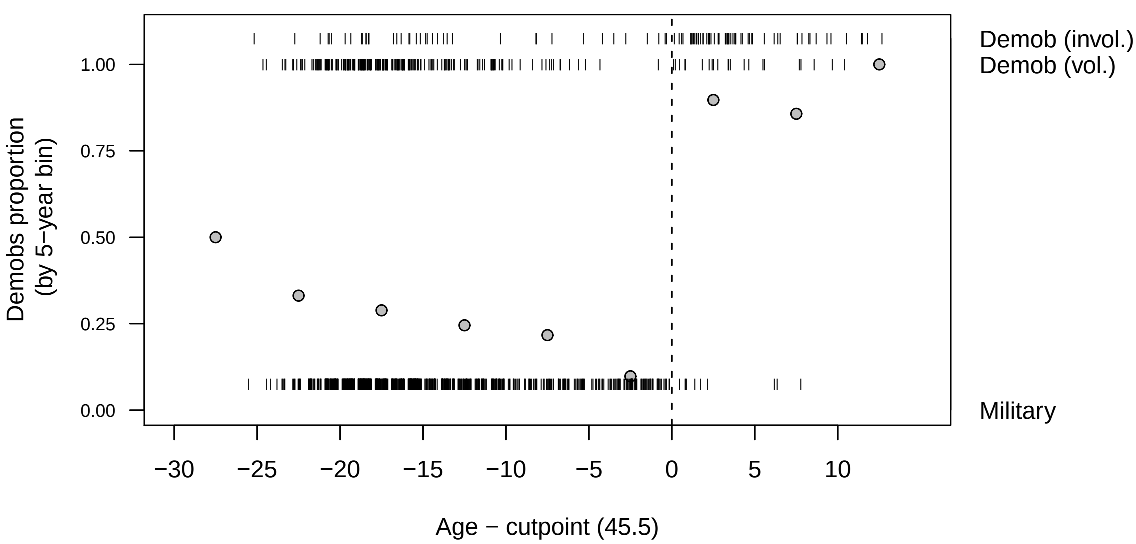


Figure 2.3: **Relationship between the prejudicial behavior and salience measures)**

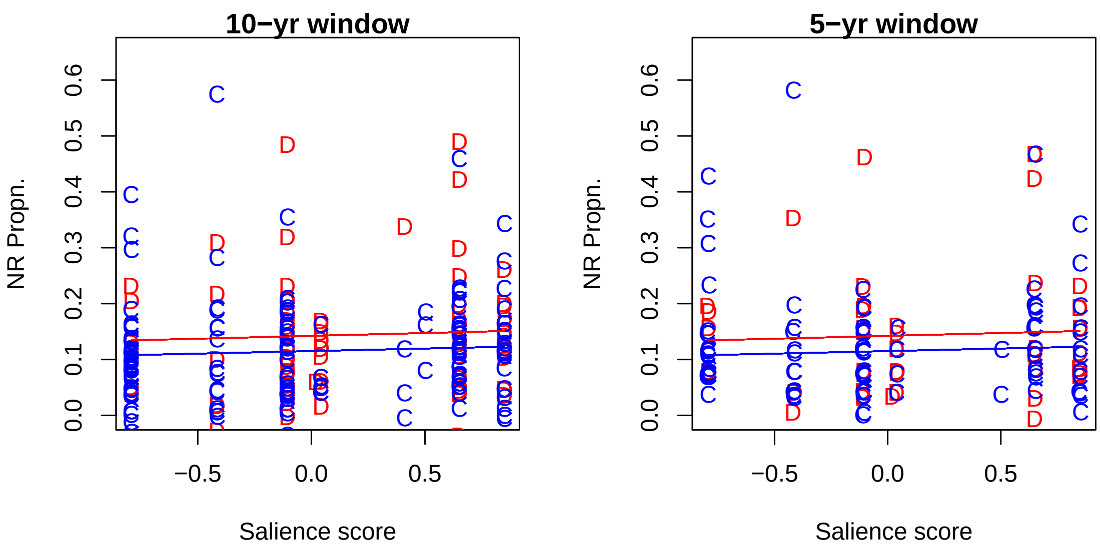


Figure 2.4: **Effects on prejudicial behavior (plots of reduced form results on the raw data)**

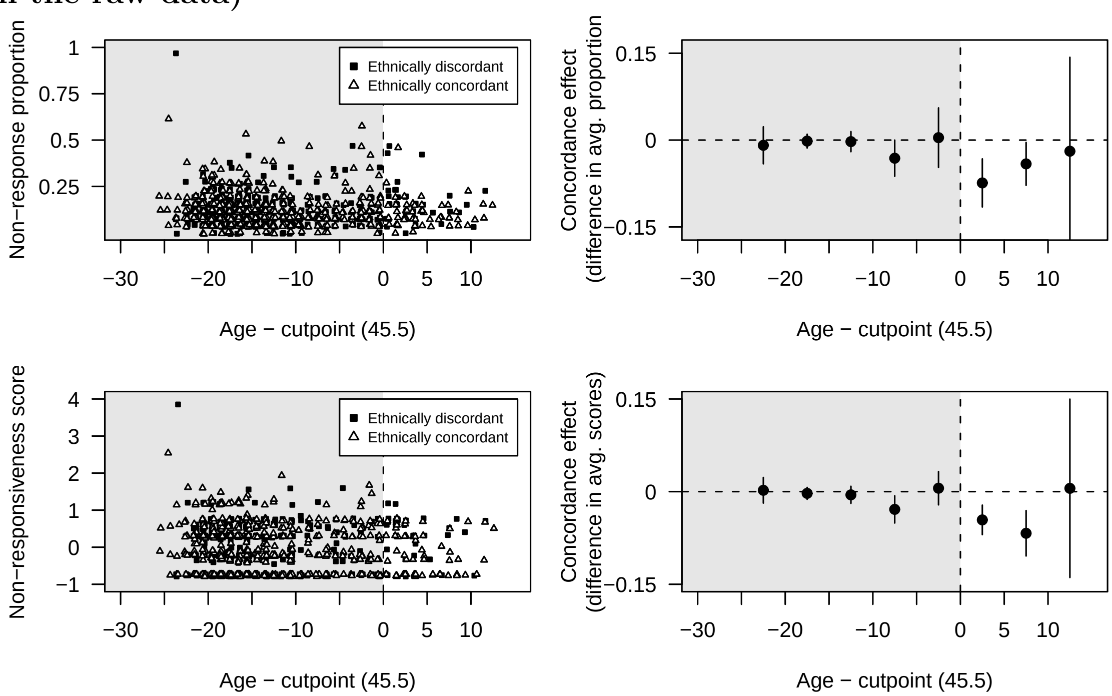


Figure 2.5: **Effects on expressions of ethnic salience (plots of reduced form results on the raw data)**

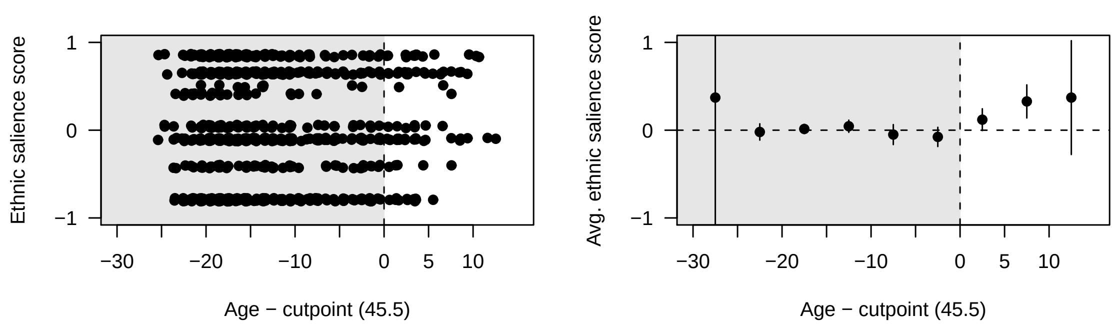
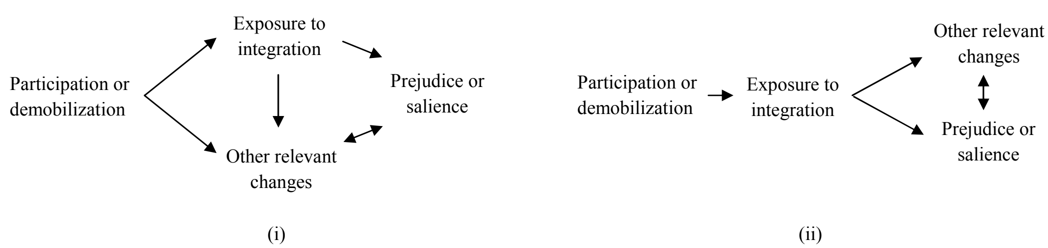


Figure 2.6: **Alternative mechanisms relating participation/demobilization to prejudice and salience**



Graph (i) on the left shows a causal pathway where effects other than exposure to quota-based integration may affect prejudice or salience. Graph (ii) on the right shows a causal pathway where exposure to integration is a primary effect of participation, and prejudice, salience, and other relevant changes follow from that.

Appendix of Ancillary Tables and Figures

Questions used for non-responsiveness score

Question id	Question
dm26	In 1993, which political party or movement did you support?
dm27	Which political party do you currently support?
cm5	Where you live now, how wealthy are you compared to others?
cm6	…compared to other Hutus?
cm7	…compared to other Tutsis?
cm8	Before the war in your community, how wealthy was your faimly compared to others?
cm9	…compared to other Hutus?
cm10	…compared to other Tutsis?
deathroster1	Did anyone in your immediate family die during the war?
pf30	What was the main cause that the CNDD-FDD was fighting for?
pf31	…that the FNL was fighting for?
pf32	…that the FAB was fighting for?
tr3	Did you witness civilians being killed?
tr6	How many members of your family were killed in the war?
tr7	…friends were killed in the war?
uo1	What was the name of your fighting unit?
lb2intara	Where did your last combat engagement take place?
re1	How do people in your community look upon former rebels?
re2	…former FAB?
re4	Some people say that former combatants who killed civil populations or who raped women should not be accepted in their families in any case and they should be punished. Some others say that they should be accepted and what happened should be forgotten. A third group says that they could be accepted if they beg for forgiveness. Which of the three groups do you support?
cp3	Comparing with the situation before the war, do you think you have more, fewer or the same political rights?
cp6	If there are persons who were rich before the war due to ethnical, regional or political exclusion, do you think that: [read the two options] 1. The government should seize their properties in order to use them for public interests. -- or -- 2. The government does not have the right to take those properties, as is the case for any other person?
cp8	Which one of the following statements do you support? 1. The government should ensure equal access to higher education as well to government jobs for all ethnic groups according to the proportions of the populations in the country -- or -- 2. The government should not consider ethnicity when recruiting for jobs or higher education institutions?
qe1	[To enumerator] Did the respondent seem distracted?
qe2	[To enumerator] Was the respondent readily willing to answer questions?
qe3	[To enumerator] Were you uncomfortable with this respondent?

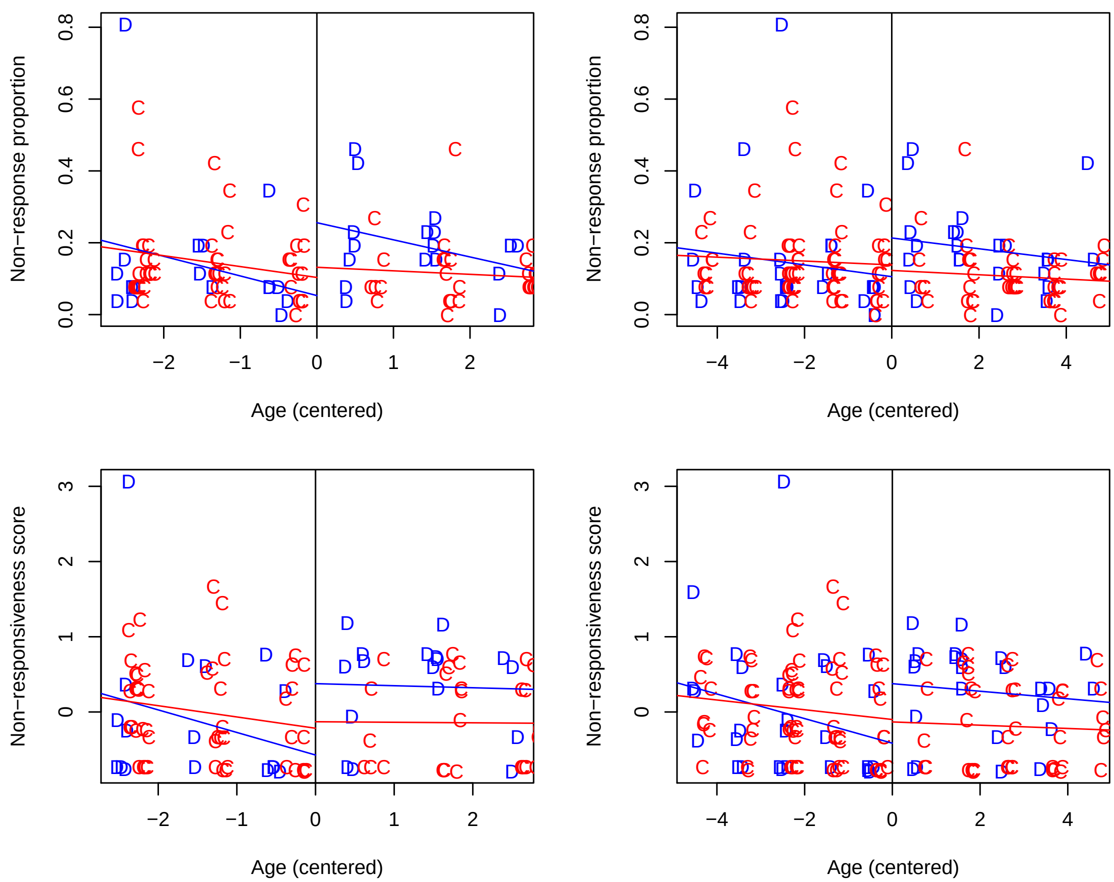
Summary statistics

Table 2.4: Summary statistics

	Full sample (N=1086)			**Within 10-yr window** (N=264)			**Within 5-yr window** (N=160)		
Variable	**Mean**	**S.D.**	**N**	**Mean**	**S.D.**	**N**	**Mean**	**S.D.**	**N**
Hutu	0.17	0.38	1086	0.07	0.25	264	0.07	0.25	160
Age (centered at 45.5)	-13.17	7.64	1086	-1.96	4.81	264	-0.35	2.64	160
Non-response proprtn.	0.12	0.09	1086	0.13	0.10	264	0.14	0.12	160
Non-response factor scr.	-0.08	0.64	1086	-0.07	0.65	264	0.00	0.67	160
Eth. salience factor scr.	0.02	0.61	980	0.01	0.59	238	0.01	0.59	142
Below cutpoint	0.91	0.28	1086	0.66	0.47	264	0.57	0.50	160
Ethnic conc.	0.65	0.48	1085	0.67	0.47	264	0.68	0.47	160
Military	0.68	0.47	1086	0.60	0.49	264	0.56	0.50	160
NCO	0.72	0.45	1017	0.97	0.18	234	0.99	0.12	138
Military years	12.68	6.82	1074	21.89	7.03	256	24.24	5.59	152
Pre-war ed.	5.83	2.22	1086	7.45	2.22	264	7.40	2.19	160
Unit death rate	0.09	0.14	1002	0.09	0.17	239	0.08	0.16	143
Family death rate	0.12	0.17	1086	0.14	0.17	264	0.15	0.18	160
Economic condns. "bad"	0.50	0.50	946	0.52	0.50	235	0.54	0.50	137
Log monthly income	10.55	0.77	1029	10.57	0.84	248	10.57	0.86	147

Examining differences in baseline non-responsiveness levels

Figure 2.7: **Non-response proportions and non-responsiveness factor scores for subjects in ethnically discordant (D, blue) and concordant (C, red) pairs, with local linear regression fits**



The red lines are local linear regressions for the subjects in ethnically concordant pairs, the blue lines are for subjects in the ethnically discordant. The graphs to the left show results within the 3-year bandwidth window, and the graphs to the right show results in the 5-year bandwidth window. The points are jittered to show massing on common values.

Table 2.5: **TSLS regressions of non-response proportions and non-response factor scores, reversing coding of ethnic concordance treatment variable**

	(1)	(2)	(3)	(4)	(5)	(6)
	nrprop	nrprop	nrprop	prej	prej	prej
military	-0.11	-0.01	-0.03	-0.30	-0.07	-0.29
	(0.22)	(0.06)	(0.05)	(0.56)	(0.38)	(0.31)
militaryXethdis	-0.04	-0.07*	-0.06*	-0.55*	-0.56**	-0.40**
	(0.10)	(0.04)	(0.03)	(0.32)	(0.25)	(0.19)
ethdis	0.06	0.08***	0.07***	0.49**	0.54***	0.54***
	(0.06)	(0.03)	(0.02)	(0.19)	(0.17)	(0.13)
age_c	-0.03	-0.01	-0.03	-0.14	-0.05	-0.16
	(0.05)	(0.01)	(0.02)	(0.11)	(0.08)	(0.10)
below_cXage_c	0.01	0.01	0.02	0.08	-0.01	0.08
	(0.02)	(0.01)	(0.02)	(0.07)	(0.07)	(0.10)
age_csq			0.00			0.01
			(0.00)			(0.01)
below_cXage_csq			-0.00*			-0.03**
			(0.00)			(0.01)
_cons	0.19	0.14***	0.16***	0.14	-0.06	0.11
	(0.14)	(0.04)	(0.04)	(0.33)	(0.24)	(0.22)
N	110	160	264	140	160	264

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (1) uses an Imbens-Kalyanamaran optimal bandwidth of 3 years, and models (2) and (3) use 5-year and 10-year bandwidths, respectively. Model (4) uses an Imbens-Kalyanamaran optimal bandwidth of 4 years, and models (5) and (6) use 5-year and 10-year bandwidths, respectively.

TSLS estimates with enumerator fixed effects

Table 2.6: **TSLS regressions of non-response proportions and non-response factor scores, with enumerator fixed effects**

	(1)	(2)	(3)	(4)	(5)	(6)
	nrprop	nrprop	nrprop	prej	prej	prej
military	-0.29*	-0.12*	-0.13*	-0.93**	-0.61*	-0.72*
	(0.17)	(0.07)	(0.07)	(0.47)	(0.35)	(0.40)
militaryXethconc	0.05	0.07	0.05	0.52	0.46	0.41**
	(0.09)	(0.06)	(0.04)	(0.32)	(0.32)	(0.18)
ethconc	-0.03	-0.09	-0.06*	-0.43	-0.59**	-0.58***
	(0.08)	(0.06)	(0.03)	(0.30)	(0.27)	(0.17)
age_c	-0.07	-0.02	-0.04	-0.13	-0.07	-0.13
	(0.05)	(0.02)	(0.03)	(0.10)	(0.07)	(0.12)
below_cXage_c	0.00	0.02	0.03	-0.00	0.02	0.03
	(0.04)	(0.02)	(0.03)	(0.07)	(0.09)	(0.13)
age_csq			0.00			0.01
			(0.00)			(0.01)
below_cXage_csq			-0.00*			-0.02
			(0.00)			(0.01)
_cons	0.30***	0.28***	0.29***	0.94**	0.93**	1.00***
	(0.09)	(0.08)	(0.07)	(0.37)	(0.37)	(0.31)
N	108	157	261	138	157	261

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (1) uses an Imbens-Kalyanamaran optimal bandwidth of 3 years, and models (2) and (3) use 5-year and 10-year bandwidths, respectively. Model (4) uses an Imbens-Kalyanamaran optimal bandwidth of 4 years, and models (5) and (6) use 5-year and 10-year bandwidths, respectively.

Table 2.7: **TSLS regressions of ethnic salience factor scores, with enumerator fixed effects**

	(1)	(2)	(3)
	salience	salience	salience
military	0.29	0.09	-0.18
	(0.39)	(0.31)	(0.15)
age_c	0.14	0.09	0.03
	(0.09)	(0.07)	(0.02)
below_cXage_c	-0.04	-0.05	-0.04*
	(0.09)	(0.08)	(0.03)
_cons	0.38	0.41	0.53***
	(0.32)	(0.31)	(0.16)
N	123	139	235

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (1) uses an Imbens-Kalyanamaran optimal bandwidth of 4 years, and models (2) and (3) use 5-year and 10-year bandwidths, respectively.

Assessing the exclusion restriction for effects of quotas

Table 2.8: **TSLS regressions of subjective economic well-being**

	(1)	(2)	(3)
	ecperbad	ecperbad	ecperbad
military	-0.19	-0.09	-0.06
	(0.33)	(0.26)	(0.16)
age_c	0.05	0.08*	0.01
	(0.07)	(0.05)	(0.02)
below_cXage_c	-0.12	-0.13**	-0.01
	(0.09)	(0.06)	(0.02)
ethconc	0.12	0.19**	0.14**
	(0.11)	(0.09)	(0.07)
_cons	0.40**	0.30**	0.45***
	(0.19)	(0.15)	(0.10)
N	120	137	235

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (1) uses an Imbens-Kalyanamaran optimal bandwidth of 4 years, and models (2) and (3) use 5-year and 10-year bandwidths, respectively.

Table 2.9: **TSLS regressions of log(income+1)**

	(1)	(2)
	linc	linc
military	0.42	0.22
	(0.45)	(0.45)
age_c	-0.04	-0.09
	(0.09)	(0.12)
below_cXage_c	0.04	0.06
	(0.08)	(0.13)
age_csq		0.00
		(0.01)
below_cXage_csq		-0.00
		(0.02)
_cons	10.38***	10.51***
	(0.29)	(0.30)
N	147	248

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model (1) uses an Imbens-Kalyanamaran optimal bandwidth of 5 years, and model (2) uses a 10-year bandwidth.

Chapter 3

Who wants to forgive and forget? Transitional justice preferences in post-war Burundi

Summary

Original survey data from Burundi are used to study transitional justice preferences during a democratic transition to peace after civil war. The study focuses on preferences toward *punishing* those who perpetrated human rights violations during war and *pursuing a commonly accepted truth* about such violations. The study accomplishes three things. First, it employs a specially-devised questioning method to measure the general extent of support for punishment and truth-seeking, with minimal social desirability bias. It finds that most Burundians take a moderate position on punishment and truth-seeking, contrary to claims in the advocacy literature. Second, it assesses whether support may be motivated by political tendencies, insecurity, or lack of knowledge. It finds support for all three, although the evidence speaks most clearly in favor of political motivations. Finally, the study uses a persuasion experiment to examine how responsive are people to attempts at persuading them to change their views. The results show that people are not very responsive, and

that if there is any reaction, it is that people become more resolute in their position. The results suggest that the international community should proceed with caution in pursuing transitional justice measures, paying due attention to how transitional justice interventions may affect hard-won political compromises.

3.1 Introduction

What informs public support or skepticism toward transitional justice after civil war? What are the implications for post-conflict transitional justice initiatives? To address these questions, I use original survey data from Burundi. I focus on preferences toward *punishing* those who perpetrated human rights violations during war and *pursuing a commonly accepted truth* about such violations. I examine what a specially-devised questioning method that minimizes the potential for social desirability bias reveals about general levels of support for punishment and truth-seeking. I assess whether variation in support may be motivated by political tendencies, insecurity, or lack of knowledge. Finally, I use a persuasion experiment to study how responsive are people to attempts at persuading them to change their views. Altogether, I find that support for punishment and truth-seeking is probably lower than what the advocacy literature suggests, that political tendencies likely influence expressions of support, and simple attempts at persuasion may backfire.

Transitional justice mechanisms are the subject of intense debate. Some have suggested that aggressive pursuit of transitional justice via the International Criminal Court has unduly burdened peace processes in Northern Uganda and the Democratic Republic of Congo.[1] Controversy surrounds the use of amnesties as incentives in many peace process (Freeman, 2010). In Liberia, controversy surrounds the work of the national truth commission. When the Commission's work was done, it issued recommendations that included barring standing president Ellen Johnson-Sirleaf from holding office again after her current term, causing some to conclude that the Commission had damaged the consensus needed to strengthen the rule of law (Gettleman, 2009). In Burundi, the case under consideration in this paper, there has been considerable controversy too. During the peace process, United Nations representatives, non-governmental organizations, and representatives of minority-Tutsi-led opposition

[1] See the discussion in Thom et al (2008).

parties have debated vigorously with the post-war government led by the former rebel National Council for the Defense of Democracy - Forces for the Defense of Democracy (CNDD-FDD, by their French acronym) over the establishment of a tribunal. The latter has insisted that a public "pardoning" process precede any tribunal, with amnesty from the tribunal being offered to those who confess and receive pardon. International observers favoring tribunals have suggested that the CNDD-FDD has been self-serving in halting an important step toward overcoming that country's cycle of Hutu-Tutsi violence. The CNDD-FDD have rebutted such claims, suggesting that the international community is being manipulated by an opposition of displaced elites intent on shackling the popular CNDD-FDD.

The transition in Burundi, like many other contemporary transitions, is one in which belligerent parties ended their fighting under an agreement to put a democratic process into place. Such processes may fail to produce "genuine" democracies. Nonetheless, the democratic transition model is the standard that the United Nations, major western donors, and regional organizations (e.g., the European Union and African Union) have used in assisting recent war-to-peace transitions, for better or worse (Newman et al, 2009). This study contributes to our working knowledge on designing such processes, focusing on transitional justice. Arguably, preferences of citizens ought to inform judgments of how transitional justice measures should be incorporated. Victims voices may deserve the greatest weight, but opinions from non-victims also help to obtain a full picture of likely consequences of transitional justice measures. Consultation with citizens helps to ensure that the appropriate political or psychological factors are taken into account. Also, it is the citizens, not international policy-makers, who bear the costs if things go awry (Snyder and Vinjamuri, 2003). At the same time, we might question whether a genuine canvassing of public preferences is possible after civil war. Many people may be under the duress of lingering fear. Segments of the public may misunderstand what transitional justice processes entail. If so, then citizens may not be able to express their "real" preferences over transitional justice alternatives.

Burundi provides a useful case for studying these questions. The timing of the study was such that transitional justice questions were still open. No formal processes had been established, although there calls to do so. Thus, the study offers a glimpse at "unadultered" transitional justice preferences in an immediate post-war context. Burundi is an important case for students of transitions from war to peace. Structurally, Burundi's Hutu-Tutsi ethnic structure is part of the class of "ranked ethnic systems" that have provided the setting especially difficult violent conflicts (Horowitz, 1985; Wimmer, 2006). Privilege afforded to segments of the minority Tutsi and the severe constraints to mobility for majority Hutu provide a ready narrative for ethnic mobilization that has colored Burundi's, and neighboring Rwanda's, violent post-independence history. Lessons from this case may be applicable to other contexts facing similar challenges.

I begin with a discussion of theories about why people may prefer punishment over forgiveness or truth-seeking over "just letting things go" after civil war. I discuss findings from the current literature, which consists mainly of surveys meant to facilitate transitional justice interventions. I explain how these may present a biased view. I follow by discussing my methodology for minimizing such bias, and I describe my survey data from Burundi. Next, I present results on general levels of support for punishment and truth-seeking. I then present findings on how political tendencies, insecurity, and lack of knowledge may affect the aggressiveness of demands for punishment or truth. Following that, I present a persuasion experiment that studied how responsive subjects' preferences might be to attempts at persuasion. The conclusion draws out implications and proposes areas for future work.

3.2 Public expressions of transitional justice preferences

Transitional justice mechanisms are common in post-civil war contexts today, although not ubiquitous (Thoms et al, 2008). A call for such mechanisms, including calls by non-governmental organizations (NGOs) like Amnesty International, Human Rights Watch, and the International Center for Transitional Justice, accompanies nearly all current peace processes. This is especially so for transitions based on negotiated settlements, most of which involve international mediation and assistance.

Transitional justice mechanisms may seek to establish a formally recognized truth about abuses. They may seek to punish human rights abusers by executing them, jailing them, or curtailing their rights to hold public offices. Transitional justice mechanisms vary in the vigor with which they pursue punishment or truth. They may also include measures for reparation.

Such mechanisms find theoretical justification in the propositions articulated in the United Nations (2004) policy document on *The Rule of Law and Transitional Justice in Conflict and Post-Conflict Societies*. This is a focal document for international transitional justice practitioners. Key propositions are as follows. First, punishment of abusers contributes to a norm of accountability that may prevent future abuses and provide a lesson on how scores can be settled through a legal process. Second, holding the main perpetrators accountable helps to separate them from larger groups (e.g. ethnic groups) to which they belong, opening space for moderate leaders to emerge and helping to end inter-group resentments. Third, the establishment of a formal truth makes legitimate punishment and reparations possible. Fourth, the establishment of such truth may also have a reconciliatory effect by lessening the mistrust of the abused about whether their plight will be recognized and creating an opportunity for forgiveness. Thom et al (2008) and Lie at al (2007) discussed theorized effects of transitional justice mechanisms from the social science, psychology, and legal literatures.

Pursuing transitional justice involves normative trade-offs (Bass, 2001; Kaminski et al, 2006; Lie at al, 2007; Pankhurst, 1999; Rotberg, 2000; Snyder and Vinjamuri, 2003; Thom et al, 2008). Opportunities to learn the truth may have to be traded off against opportunities to punish. Opportunities to contribute to norms of accountability may have to be traded off against opportunities to get belligerents to buy into a given peace process. Opportunities to establish common truths may have to be traded off against risks of exciting tensions. The sources of these trade-offs are clear. The greater the threat of punishment, the less likely will perpetrators compromise positions of power in peace negotiations, and the less likely are perpetrators and their associates to volunteer important information about what happened. In seeking a "formal" truth, one may come up against events that are highly ambiguous, and delineations of perpetrator and victim may be inciting.

In a democratic context, the citizens who hold some claim to justice due to exposure to abuses should expect to have their voices taken into account in managing these dilemmas. Among this section of the public, though, there are likely to be constituencies for and against the robust punishment or truth-seeking. Making matters more complicated, there are reasons to suspect that public expressions in *favor* of transitional justice processes may not reflect privately held beliefs about what is just. For example, support for punishment and truth-seeking may be based on agendas to undermine the political standing of those that would be targeted by such mechanisms. The ostensible *objectivity* of transitional justice processes does not imply that they are politically *neutral* in their effects. A conviction of a leader will change the political balance in a manner the disfavors the group represented by that leader.

Going in the opposite direction, expressed skepticism toward punishment or truth processes may also be based on things other than justice perceptions. Resignation of one's "right" to pursue punishment or truth may be due to a sense of duress due to fear of imminent danger. This tendency may be reinforced by people's sense that by voicing demands for transitional justice, they may be creating an unwelcome disturbance. If so, the lack of demand for transitional justice measures is more the result of a "spiral of silence" (Noelle-Neumann, 1984) or "preference falsification" (Kuran, 1998) than the expression of true desires. Then, transitional justice ideals may be worthy, but to be workable they need to be married to a strategy for ensuring security. Another "illegitimate" reason that someone may not express a demand for transitional justice could be a lack of understanding about what might be possible (Sudharshan, 2003). If so, then there is no reason to question transitional justice ideals per se; rather, there is the need to educate the public more.

Another possibility is that individuals tend to have good reason to prefer not to pursue punishment or truth. Skepticism may be a "pragmatic" response to "political realities" (Snyder and Vinjamuri, 2003). Perhaps there are more remote concerns about reigniting conflict—not concerns expressed under duress as much as concerns expressed under acceptable conditions that are not worth undermining. Or, political and institutional changes brought about by the war may sufficiently endow former victims with a sense of empowerment and rights protection.[2] If so, those who were formerly vulnerable to abuse may prefer to move on. There may even be costs: pursuing punishment or truth may induce productive members of society to flee. Or, the ambiguity of past events may be such that the attempts to delineate perpetrators and victims may be tenuous at best, and may contribute on balance to undermining cooperation.

In interpreting mass data on public attitudes toward transitional justice, one cannot necessarily use responses of "for" or "against" as straightforward, context-free measures of privately held preferences. Some attention to *why* people express these positions is in order. If support constituencies are based primarily on motivations of

[2]Theidon (2006) discovered that a new sense of empowerment motivated a preference for reconciliation and even forgiveness over punishment among communities in Ayaucho in in post-war Peru.

gaining political advantage, transitional justice professionals should proceed very cautiously so as to be sure not to inadvertently undermine a hard-won political balance. Skepticism due to fear or lack of understanding would lead us to propose strategies for security-enhancement and education in the service of transitional justice ideals. If such illegitimate sources of skepticism are not plausible, then the onus is on transitional justice practitioners to justify transitional justice ideals before a skeptical public. The empirical analysis below examines these propositions.

3.3 Related literature

Social scientists have only recently begun to study public attitudes toward transitional justice in post-conflict settings.[3] This is remarkable given the accelerated rise in resources committed to post-conflict transitional justice since the early 1990s (Thom et al, 2008). But it is also quite understandable given the logistical difficulty of sample surveys in post-conflict settings. Because of this difficulty, nearly all quantitative research in this domain has been done by researchers from a common network that links United Nations agencies, particularly offices of the United Nations Development Programme, with non-governmental organizations advocating transitional justice, including the International Center for Transitional Justice, the BBC World Trust, Search for Common Ground, and the Human Rights Center at University of California, Berkeley. Some the research is of very high quality, being submitted to rigorous peer review (e.g. Pham et al, 2004). But there is a potential conflict of interest: organizations advocating-for and depending-on the promotion of transitional justice measures have incentives to paint a picture favorable to such measures. Other work in this domain includes a survey conducted by the Afghanistan Independent Human

[3]Complementing the studies on public attitudes is a literature that studies the apparent *impact* of transitional justice mechanisms. This literature has been recently reviewed in Thom et al (2008). Questions about effectiveness are not the subject of this paper.

Rights Commission (AIHRC, 2005) as part of national consultations on transitional justice.

Among available studies, only Pham et al (2004) study systematically potential motivations for taking more or less aggressive positions toward pursuing transitional justice. Their key explanatory variables of interest are associated with exposure to wartime traumas. They propose that trauma may reduce one's sense of "self-efficacy" (Bandura, 1977) and thereby induce people to resign rights to transitional justice. They propose that this effect is mediated by whether one is psychologically resilient or succumbs to post-traumatic stress disorder. They find mixed evidence for the hypothesis. The other studies provide either basic percentages of respondents' expression one or another opinion on transitional justice measures, or at most, cross tabulations based on region, education, victim status, or gender. Table 3.2 summarizes these studies. I report modal response percentages for questions relevant to the current discussion.

The studies tend to find support for accountability measures, although the salience of transitional justice appears to be low: never was the pursuit of justice a priority over basic needs and development. The nature of the support for "accountability" is sometimes difficult to discern. In an impressive repeated cross-section study, Pham et al (2005; 2007) show that respondents in war-affected districts of Northern Uganda take opinions that would seem to contradict each other. While 66% of 2005 respondents indicated that LRA offenders should be punished, 65% reported that they would accept amnestied LRA leaders if they were to return home. The percentage of respondents indicating that LRA offenders should be tried dropped to 32% in 2007, likely due to a perception that the ICC's pursuit of LRA leader Joseph Kony derailed the peace process (Pham et al, 2007:47). In the other cases, we see more ardent support for punishment. Responses about truth-seeking tend to be more consistently in favor.

There are a few reasons to be cautious about the findings of these surveys. First, we need to be careful about generalizing. The nature of the sample universes differ from study to study, with some attempting to represent entire national populations and others limiting themselves to groups most affected by wartime violence. Second, none of the studies address in depth the types of motivations discussed above. A few of the studies report some suggestive findings. For example, Pham et al (2005) find that in Northern Uganda, respondents from non-Acholi war-affected districts were twice as likely to prefer "peace with trials" over "peace with amnesty" as compared to those from war-affected Acholi districts. UNDP (2007) finds that in Kosovo, Kosovar Albanians are twice as likely to support the international tribunal over Kosovar Serbs. In both of these cases, ethnic markers define major political cleavages.

Third, the manner in which the issues are posed in the questionnaires may bias results in favor of pro-transitional justice responses. The questions tend to be posed in a one-sided "yes" or "no" manner, with respondents typically being asked whether they would like to have a tribunal, truth commission, or some other mechanism put in place. Respondents are not asked to choose between options presented in a more neutral manner. The psychology of survey response suggests that this may be leading. Survey respondents will respond to cues about what response is desired by those issuing the survey. Thus, respondents may be inclined to say "yes" for reasons of social desirability (Converse and Presser, 1986). An interesting exception was in Pham et al's (2005) study in Northern Uganda, when respondents were given the choice between favoring "peace with amnesty" or "peace with trials." On that question responses were nearly split down the middle, with 56% of respondents in Acholi districts expressing a preference for "peace with amnesty" and only 39% expressing the same in non-Acholi districts.

Finally, it is hard to say from this research how deeply held are these attitudes and how responsive they might be to events or persuasion. The Pham et al (2007) study in Northern Uganda, which updates their 2005 study, suggests that opinions may be quite responsive to events. Changes in aggregate opinions between the two surveys could reflect reactions to the controversy over Kony and the ICC. It would be useful to study such responsiveness more directly.

With survey evidence that I present below, I attempt to overcome a number of these shortcomings. I use a questioning method designed to minimize social desirability bias. I make a focused effort to study the motivations discussed above. Finally, I use a deliberation experiment to assess the responsiveness of people's attitudes to attempts to persuade them to change their mind.

[Table3.2 here.]

3.4 Case context

Burundi is a small, impoverished, and land-locked country of approximately 8 million people (ca. 2010) in central Africa. Like Rwanda to the north, Burundian society is marked by a caste-like stratification that has historically privileged a Tutsi minority relative to majority Hutu and a very small third group, the Twa. Like in Rwanda, Burundians have struggled to escape a conflict pitting custodians of this "ranked ethnic system" (Horowitz 1985; Lemarchand, 1970) against those ostensibly seeking to remove barriers to Hutu mobility. The country's history is marked by bouts of genocidal violence and barbarous repression. Most notable are the events of 1972, when a Hutu insurrection escalated to involve massacres of Tutsis, mostly in the southern part of the country. This triggered a massive crackdown by the Tutsi-dominated army, which went beyond restoring order and sought to prevent future uprisings by "decapitating" Hutu society. The estimated number killed in that violence—mostly Hutu, it is thought—is 150,000-200,000, with massive outflows of Hutus into neighbouring Rwanda and Tanzania (United Nations, 1996). The ensuing decades involved increasing concentration of authority in hands of the southern-Tutsi, military elite and short bouts of insurrectionist violence.

A period of liberalization in the early 1990s led to elections in 1993. These resulted in the triumph of a party that represented the aspirations of a long-oppressed Hutu majority. But under still-mysterious circumstances, members of the southern- and Tutsi-dominated army led a bungled coup attempt in October 1993 that involved the assassination of the recently-elected Hutu president. The event triggered massive violence throughout the country, and the ensuing ferment gave rise to a formidable rebellion. The fighting between the government and rebel forces was episodic over the ensuing decade. It touched most of the country, resulting in an estimated 300,000 deaths. Major hostilities ended when the largest rebel group, the CNDD-FDD, signed onto the peace process in Pretoria in 2003. In the 2005 elections, the CNDD-FDD won an outright majority of national assembly seats (59% of 118 seats) and communal councilor posts (55% of 3,225 posts) in the 2005 elections. The CNDD-FDD's political head, Pierre Nkurunziza, was elected to be president. Thus, the war resulted in a *near revolution* in the institutionalized political context relative to the pre-war status quo.

The survey interviews were conducted in Burundi in June-August 2007, two years after the elections. A small splinter from another rebel faction, the Hutu "liberationist" *Front national de liberation - Parti pour la Libration du Peuple Hutu* (FNL-PALIPEHUTU), remained at large, although this did not impede the survey.

At the time of field work, no formal transitional justice processes had been initiated. The United Nations had put forward considerations for transitional justice in Burundi as early as 1996 in a special commission report on the 1993 violence (United Nations, 1996). The commission found reason to believe that organized, genocidal Hutu-on-Tutsi violence had taken place. The commission also recognized that this episode was part of a long cycle of violence, and that any measures should take this into account. With the war raging, there was no possibility of action being taken. The conversation about truth and justice processes was confined largely to elites in Burundi's capital and donor capitals. Transitional justice measures entered the discourse again during the talks that led to the 2000 Arusha Accords. The Accords called for establishing a truth commission and a "special chamber" to try those accused of genocide. However, the CNDD-FDD were not party to the Accords. When the CNDD-FDD signed onto the peace process in 2003, its leadership suggested that questions of truth commissions and special chambers would have to be revisited after elections. As discussed in the introduction, the issue was subject to ongoing debate between the CNDD-FDD on one side and the UN and opposition parties on the other. At the time of fieldwork, the issue of whether and how to implement transitional justice measures for Burundi remained a wide open question.

3.5 Methodology

3.5.1 Question design

[Figure 3.1 here.]

Questions about transitional justice topics are normatively loaded, and thus likely to be subject to social desirability biases unless careful attention is paid to minimizing it. To reduce the potential for such bias, I designed questions and a question-delivery approach that would create a perception of *equal social legitimacy* to alternative viewpoints. The method has two elements: wording of the choices and gestures that accompany the verbal delivery. These elements are shown in Figure 3.1. The wording of the questions that about preferences for *punishment* and *truth-seeking* are at the bottom of the figure. The punishment question asks the respondent about his or her preferences over three possibilities: punishing human rights violators, forgiving them unconditionally, or forgiving them only conditional on their admitting having done something wrong. Each choice is read as being favored by "some people," cuing the respondent to appreciate that he or she would not be alone in taking any of the options. To enhance the cue, enumerators were trained to use specific physical gestures, as shown in Figure 3.1. The gestures allow the respondent to visualize the different groups associated with each respondent. A gestured delivery such as this also bears resemblance to everyday conversational practices among Burundians. The truth-seeking question is slightly different in its phrasing. Because truth-seeking in the Burundi context typically refers to the events of 1993 and 1972, the question asks about "what happened before the war." The respondent is asked to express a preference for learning the truth versus "forgetting" the past. Both of the choices have the same pre-amble, "In order to achieve peace and reconciliation...", cuing the respondent to see either as a legitimate option for promoting these goals. The question is delivered with accompanying gestures to cue the existence of two equally legitimate options. Enumerators were trained extensively on the questioning procedure, including the precise wording and the use of gestures. The question design takes seriously the fact that a survey interview is a scripted social interaction.

Other data used below include demographic and geographic information. This information was collected in a more straightforward manner, given the non-vulnerability to social desirability bias.

3.5.2 Persuasion experiment design

[Figure 3.2 here.]

I also embedded in the survey a persuasion experiment associated with the question on punishment. The experiment was similar in design to that used by Jackman and Sniderman (2006) to study the effects of deliberative discussion on attitudes toward labor laws in France. The design of the experiment is displayed in Figure 3.2. Each respondent was asked the punishment question, as described above. Then, depending on the answer that was given, each respondent was randomly assigned to receive either a "vacant" or a "content-laden" counterargument. To prevent error in the implementation of the randomization process, the questionnaires were pre-printed with either the content-laden or vacant counter-arguments, and then randomly shuffled into the stacks given to the enumerators. Figure 3.2 shows that distributions are mostly even across the treatment conditions. By random chance, it appears that some asymmetry emerged in the shuffling of questionnaires for the counter-argument conditions for choice 3.

In order to measure the possible extent of preference falsification, I sought to measure the extent to which the counter-arguments would cause respondents to change their response. Furthermore, in order to assess the possible effectiveness of difference types of persuasion, I sought to determine whether persuasion would be more effective in moving people toward more or less aggressive positions. Once the counter-argument was delivered, the respondent was asked again about what choice he or she preferred. The vacant counter-argument was the same for all choices. No matter what choice the respondent gave, the enumerator followed with,

> *Vacant counter-argument*: However this can lead to some difficulties. Then I would like to ask you again. Do you think it is good to (1) punish them, (2) accept them when they come back, (3) ask them to beg for forgiveness?

The content-laden counter-arguments were specific to the choices that the respondent made initially. If the respondent chose option 1, to "punish", then the enumerator would say,

> *Content-laden counter-argument 1*: But there are people who say that both sides have committed many crimes during the war, thus it is the time for people to forgive so that we can progress. So I would like to ask you again. Do you think it is good to: (1) punish them, (2) accept them when they come back, (3) ask them to beg for forgiveness?

If the respondent chose option 2, to "forgive" unconditionally, then the enumerator would say,

> *Content-laden counter-argument 2*:But if we ignore what happened people could be angry and take revenge. So I would like to ask you again. Do

> you think it is good to: (1) punish them, (2) accept them when they come back, (3) ask them to beg for forgiveness?

And finally, if the respondent chose option 3, to "forgive" only conditionally, then the enumerator would say,

> *Content-laden counter-argument 3*:But there are some people who think that justice is not necessary, while others assert that both sides have committed many crimes and that it is time for reconciliation. So I would like to ask you again. Do you think it is good to: (1) punish them, (2) accept them when they come back, (3) ask them to beg for forgiveness?

The different counter-arguments allow us to test the extent to which subjects might be induced to taking a more or less aggressive position. The vacant counter-arguments provide a control condition, to ensure that we do not confound persuasion with intimidation from merely receiving a counter-argument. To the extent that preference falsification is taking place in the initial responses, the experiment will help to reveal whether the falsification is masking more aggressive or more forgiving preferences.

3.5.3 Sample and post-stratification

The data are drawn from the multi-purpose survey of *Wartime and Postconflict Experiences in Burundi*. This survey was designed to serve multiple research purposes, including studies on the effects of economic conditions on participation in revolt and the impacts of security sector reform and ex-combatant reintegration programs.[4] A self-weighting sample representative of the population would be most useful for the goals of the current paper. However, such a sample would not maximize power, for a given sample size, for the other studies. Thus, the sample that we drew was based on

[4] Details on the survey and these various studies can be found at http://www.columbia.edu/~cds81/burundisurvey/

an attempt to manage these trade-offs. The sample was drawn from strata consisting of civilians, demobilized combatants, and active members of the security forces. This paper only looks at respondents from the civilian stratum, excluding any consideration of ex-combatants or current security forces members. The population of the latter are but a small fraction of the total Burundian population (less than 2%). Therefore any biases due to their exclusion should be negligible. A key compromise in our sampling plan was to set a civilian male-to-female sampling ratio of 4 to 1. The other studies were concerned mostly with those who participated in rebellion or army, and these were almost exclusively men. Thus, we needed a rich civilian male sample to use as our comparison group. In the event, due to non-response and random error, the male-to-female sampling ratio was a bit higher than 4-to-1, as shown in Table 3.3.

Geographically, the sample was chosen through a multistage process. Within each of Burundi's 17 provinces, half of the province's communes were selected at random. Communes are Burundi's second tier administrative unit, and they contain between 15,000 to 100,000 individuals (or about 3,000 to 20,000 households). Sample sizes within each of the communes were determined on the basis of estimated population sizes (given in the *Institut Statistique et des Etudes Economiques du Burundi*'s 2006 statistical yearbook) as well as considerations of whether that commune offered special analytical leverage for one or another of the impact studies undertaken as part of the survey. Each commune in Burundi is further divided into somewhere between 5-30 collines ("hills") of approximately equal population size and containing a few hundred households each. Within each of the selected communes, we chose as sampling sites the commune's central colline plus seven other selected at random. Then, enumerators were guided to position themselves in the middle of the colline, face a randomly preselected compass direction, and to approach for interview respondents nearest to the line of sight on that compass direction.

[Figure 3.3 here.]

The manner in which commune sample targets were set is a kind of "selecting on independent variables." This poses no special problems with respect to bias when the goal is to test hypotheses based on these independent variables (King, Keohane, and Verba, 1994). However it does complicate the calculation of population-level descriptive statistics. Some kind of adjustment is needed to correct for the departures from equal probability sampling. A manner by which this can be done efficiently is weighting. One may derive adjustment weights from the sampling design or by adjusting to known population distributions (Gelman, 2007). Because our sample is not tremendously large, and because I was able to obtain good information on demographic distributions down to the commune level, I prefer to use direct adjustment to population distributions via post-stratification. The post-strata interact commune with ethnicity (Hutu or not) and gender.[5] The geographic locations of survey respondents is shown in Figure 3, and demographic features of the raw and weighted sample are shown in Table 3.3.

[Table 3.3 here.]

3.5.4 Data analysis

All estimates are computed using the post-stratification weights described above. I adjust all variance estimates with stratification at the province level and then clustering at the commune level. The data analysis was conducted using the "svy" suite in Stata version 11.

[5]Commune level population numbers for males and females are from the *Institut Statistique et des Etudes Economiques du Burundi*'s 2006 statistical yearbook. Commune level ethnicity proportions are from smoothed estimates that use results of our survey, as explained in Samii (Nd.). The weights were constructed by raking to gender and ethnicity proportions in each commune, using the "survey" package in R (Lumley, 2010).

3.6 Results

[Table 3.4 here.]

Overall preferences

The overall distribution of punishment and truth-seeking preferences is displayed in Table 3.4. The results suggest considerable moderateness in overall preferences. The modal expressed preference is for conditional forgiveness combined with a preference to "forget the past." The two preferences are correlated (p-value, 0.03). This is mostly with respect to the relationship between truth-seeking and the two less-aggressive punishment preferences: those who express a preference to seek the truth about the past are significantly more likely to express a preference for conditional versus unconditional forgiveness. The relationship does not carry through to a preference for unconditional punishment, owing to extreme rarity of expressed preference for the latter.

The picture that these results imply is different than one of mostly staunch support in the studies reported in Table 3.2. For example, the BBC World Services Trust/Search for Common Ground (BBCWST/SFCG) survey conducted less than a year after our survey (see row four of Table 3.2) suggests that 68% of respondents were recorded as expressing a preference for bringing wrongdoers to trial, and 81% were recorded as expressing an opinion that a TRC would be entirely or mostly good. Those questions were asked in a yes/no format, and they did not present a set of contrasting but equally legitimate options from which to choose. We would expect that the BBCWST/SFGC survey would generate responses that veer on the side of exaggerated support, and the results are consistent with that expectation. Of course, this is only suggestive, as the BBCWST/SFCG questions referred to formal mechanisms, and we are comparing results from two separate samples. If anything, it points to need to investigate the consequences of different questioning styles within new surveys that are fielded.

3.6.1 Assessing motivations

To study potential motivations, I use regressions to approximate the relationships between transitional justice preferences and proxy measures of political tendencies, insecurity, and access to information.

As proxies for political tendencies, I use ethnicity, region of origin, and the interaction of the two. Before the war, elites allocated resources and opportunities on the basis of ethnic and regional identity. The Southern provinces were historically privileged in terms ties to ruling elite and access to publicly financed goods in the pre-war period (Ngaruko and Nkurunziza, 2000:381-384; Jackson, 2000:3; author's own field interviews). Generally speaking, Southern Tutsis were the most privileged. Southern Hutus received some externality benefits, but they were nonetheless subject to the discrimination in schooling and access economic opportunities that Hutus throughout the country suffered. After the war, the ascendant CNDD-FDD has been considered among Burundians to seek to unravel Tutsi and Southern privilege. The outcome of the war thus assigns a status of political "losers" to non-Hutus and to those originating from the Southern provinces. For ethnicity, I use an indicator for whether one characterizes oneself as non-Hutu, thus being associated with those whose privileges may be threatened. For region, I use an indicator for whether one's home region is one the Southern provinces of Bururi, Makamba, or Rutana. The south/non-south divide cross-cuts ethnicity, at times serving as a line of fracture among the predominately Hutu rebel groups as well as the Tutsi elite (International Crisis Group, 2000). I prefer to use these identity markers as measures of political tendencies rather than endogenous behaviors such as, say, vote choice in the 2005 elections. The identity markers provide a clear perspective on whether political contestation over rights and redistribution inform transitional justice preferences.[6] In line with the discussion

[6]Ethnicity and region are highly predictive of whether a respondent indicated that they "sup-

above about motivations, I expect that both of these indicators should be associated with more aggressive transitional justice preferences. The reason is that these indicators distinguish those who have lost more as a result of the redistribution of power that occurred after the war.

My proxies for insecurity include two measures. The first is an indicator of whether one experienced wartime victimization in terms of the death of an immediate family member at the hands of the rebels or the army. The second is an indicator for whether one's home commune has been host to any major insecurity-related events in the three years prior to the survey. Such events are based on reports of violence associated with banditry, ongoing rebel activity, and state human rights abuses from 2004-2007.[7] I use pre-war home commune, rather than commune of residence at the time of fieldwork, to address relocation due to either insecurity or economic opportunities. For those living in their home commune (about 77%), this measure captures the possibility that one faces local insecurity but is somehow means constrained in one's ability to relocate. For those who have relocated, this measure captures the fact that one may not be able to return home because of ongoing insecurity. To use a measure based on insecurity in one's current commune of residence would greatly understate the level of current insecurity. My expectations for these variables are mixed. On the one hand, victimization status is expected to increase one's desire to seek either truth or punitive measures. On the other hand, on-going insecurity is expected to

---

ported" the victorious CNDD-FDD in the survey. (Asking about vote choice directly was deemed too sensitive given the context, and so we only asked about which party the respondent supported.) For non-Southern Hutu, approximately 67% are estimated to support the CNDD-FDD. Among Southern Hutu, the estimate drops to about 55%, although the difference is not statistically significant at the .10 level. For non-Southern and Southern non-Hutu, the estimated percentages are 21% and 37%, respectively.

[7] The events data were compiled from Burundian and international news sources as part of the *Wartime and Post-Conflict Experiences in Burundi* project (www.columbia.edu/~cds81/burundisurvey/).

have a "chilling effect", suppressing one's willingness to express one's preference for punishment or truth-seeking. Thus, I include both indicators as well as their interaction. The chilling effect hypothesis suggests that the interaction coefficient should be negative. In addition, a further implication of the political motivations hypothesis is that victimization by rebels should have a stronger association with aggressive preferences than victimization by the army.

Finally, as a proxy for one's knowledge of what transitional justice processes may entail, I use indicators for whether one's highest level of education is above primary school. The expectation here is that more education should be associated with more awareness about the possibilities for transitional justice measures, which may translate into more aggressive preferences. Summary statistics for all variables are in the appendix.

Regression specifications account for the causal ordering of explanatory factors. For example, since ethnicity is determined at birth, one should measure the "effect" of ethnicity with a specification that excludes post-birth variables to avoid "post-treatment bias" (King and Zeng, 2006).[8] Ethnicity is then included in specifications that measure the effects of post-birth variables (e.g., education), in order to reduce possible spuriousness. A similar logic is applied in specifying and interpreting all of the regression models. The appropriate coefficient to interpret substantively for each variable is the first one that appears as one goes from left to right in the regression tables.

I use ordered logistic regression for the punishment question, with outcomes ordered from (1) unconditional forgiveness to (3) unconditional punishment. I use a binary logistic regression for the truth-seeking question, with "seeking the truth"

[8]There is a methodological debate over whether immutable traits like ethnicity can be analyzed in terms of "causal effects" (see, e.g., Morgan and Winship, 2007). Ethnicity is used here as a proxy for political tendencies, which are manipulable if one may redefine the conditions of privilege or deprivation associated with identity.

coded as 1 and "forgetting" as 0. The regression estimates use the survey weights and variance corrections accounting for stratification and clustering. This is observational data and some of the variables are endogenous to unmeasured conditions. Thus, the results only suggest causal interpretations, rather than providing clearly identified estimates of causal effects.

[Tables 3.5 and 3.6 here.]

The regression estimates are reported in Tables 3.5 and 3.6. I report odds ratios for ease of interpretation. Odds ratios equal to 1 imply no relationship, larger than 1 imply more aggressive demands for punishment and truth, and between 0 and 1 imply less aggressive demands.

The "political motivations" hypothesis proposes that those who are non-Hutu and originating from a Southern province would have more aggressive preferences, and that victims of rebel violence would have more aggressive preference than of army violence. Again, the logic here is that ethnic and regional identity dominate politicians' and their followers' claims for how economic and political resources should be allocated, due to the legacies of Hutu and non-Southern exclusion in Burundi. Army victims are expected to have less aggressive preferences because post-war political gains by the former rebel parties has provided *political* redress. The evidence on demands for punishment are somewhat consistent with these expectations. Non-Hutu identity is associated with about 61% higher odds of choosing a more aggressive position on punishment, although contrary to expectations those with non-Hutu identity originating from the South have about 50% lower odds of choosing a more aggressive position. With respect to victimization, we see that victimization by rebels is associated with 90% higher odds of expressing a more aggressive preference, whereas there is no association with victimization by the army. With respect to truth-seeking, non-Hutu identity is associated with 86% higher odds of expressing a preference for truth-seeking, and being from a Southern province is associated with 59% higher odds, although the latter is estimated very imprecisely (p-value, 0.20). An unexpected pattern emerges with respect to victimization, however, in that rebel victimization is actually associated with 54% lower odds of expressing a preference for truth-seeking; the opposite is true for army victimization, although this too is estimated very imprecisely. Further investigation suggests that this may be an exception that proves the rule: I fit another regression (not shown) including interaction terms between non-Hutu identity and the two victimization variables. The thought was that the negative relationship was due to *self-suppression* of demands for truth among Hutu victims of rebel violence. The estimates show that this may be the case: the interaction term coefficient almost precisely canceled out the coefficient on the rebel victimization coefficient, indicating that it is almost exclusively among *Hutu* victims of rebel violence that the negative relationship is manifest. Army victimization is estimated to be associated with 37% higher odds of expressing a preference for truth-seeking, although this is imprecise (p-value, 0.33). Finally, I fit models (not shown) that included all the specifications as in Tables 3.5 and 3.6, but also included an indicator variable for whether the respondent indicated that they "supported" the victorious CNDD-FDD party. Those who indicated as such had about 30% lower odds of choosing an aggressive position on punishment and about 30% lower odds of preferring truth. These results correspond to the political motivations hypothesis. However, the statistical significance of these associations dropped considerably when the non-Hutu and Southern region indicators were included, indicating high collinearity and also, in my opinion, the primacy of political identities. Thus, the bulk of the evidence is consistent with a version of the political motivations hypothesis emphasizing ethnicity, That is, preferences for transitional justice measures appear to be based on whether such measures constrain those pursuing political agendas that disfavor one's identity group relative to helping those pursuing agendas that favor one's identity group. In Burundi, ethnic identity dominates this calculation.

The "insecurity" hypothesis proposes that on-going insecurity in one's home community is likely to suppress "rightful" demands for punishment and truth via a chilling effect on conflict victims. The evidence on this is mixed. With respect to preferences for punishment, I estimate that conditional on being a victim of rebel violence, insecurity suppresses the preference for truth-seeking by about 50%, although the estimate is highly imprecise. This chilling effect is more precisely estimated for truth-seeking: conditional on being a victim of rebel violence, insecurity reduces the odds of expressing a demand for truth by about 30%. Such effects are not evident with respect to army victimization: preferences for both punishment and truth-seeking are estimated to be higher for army victims facing insecurity. Thus, we find that the evidence is consistent with this hypothesis only for the subset of individuals having been exposed to victimization by rebels. Those for whom this combination of conditions apply constitute only an estimated 7% (s.e., 2) of the population under study. Thus, while the insecurity motivation is a possibly important reason for non-aggressiveness at the individual level, it cannot explain much of the overall non-aggressiveness that I have recorded.

The "knowledge" hypothesis proposes that aggressive of transitional justice demands should be increasing in education. This hypothesis receives no support with respect to punishment: the odds ratio is almost exactly one. However, it receives strong support with respect to truth-seeking: those with greater than primary education have more than double (2.33) the odds expressing a preference for truth-seeking. The finding may reflect the fact that transitional justice is not all or nothing: individuals may prefer a process that seeks truth but not necessarily punishment.[9]

[9]The CNDD-FDD support indicator was also collinear with the higher education variable, although when the support indicator was included, results did not change appreciably for the coefficient on the higher education variable.

3.6.2 Persuasion experiment

The persuasion experiment studies (1) how responsive are people in general to counter-arguments and (2) whether attempts at persuasion are more effective in pushing individuals toward more aggressive or less aggressive stances. The latter is especially important given the political nature of transitional justice interventions. In a democratic context, those who stand to lose politically from the advance of transitional justice mechanisms would likely try to persuade the public to resign demands for punishment or truth-seeking.[10] On the other hand, principled advocates of transitional justice processes, along with those who stand to gain politically, would pose arguments persuading people to support more aggressive positions.

[Table 3.7 here.]

The persuasion experiment was limited to preferences for punishment. Survey-weighted results are presented in Table 3.7. The tables show row proportions—that is, proportions choosing any one of the three preference options under the vacant and content-laden counter-argument conditions. Outcome tables for each of the subgroups is displayed separately. Recall respondents first placed themselves into a subgroup through their initial response to the punishment question. Then, respondents were randomly assigned to receive either the vacant or content-laden counter-argument. Overall, attempts at persuasion were usually unsuccessful. Responses remain stable at least 75% of the time across the groups and treatment conditions. Surprisingly, the effect of the content-laden counter-arguments were to make those who took positions at the extreme ends of the choice spectrum *more resolved* to maintain their position. This is evident among those who initially express a preference for unconditional forgiveness: whereas under the control condition, responses remain the same 77% of the

[10] As discussed in the context section above, this has been the case in Burundi, where the dominant and former insurgent CNDD-FDD party has issued statements suggesting that the priority should be in promoting "pardon" rather than looking into the abuses that occurred during wartime.

time, the rate increases to 83% under the persuasion condition. The difference is even greater for those whose initial response is unconditional punishment: the rate at which people maintain their initial response goes from 75% in the control condition to 88% in the treatment condition. However, in the latter case we must note that the subgroup sample is quite small. For those taking the moderate "conditional forgiveness" position, the differences are negligible.

Thus, expressed preferences are not so responsive to attempts to change them. In assessing whether persuasion tends to work more in moving people toward more or less aggressive positions, it would seem that it slightly favors aggressive positions, but not because of persuasion effects. Rather, the bias in favor of aggressive positions is due to an effect of making people more *resolute* in maintaining an aggressive position. The implication is that simple attempts at persuasion may backfire, possibility contributing to polarization. Any attempt at a deliberation or sensitization intervention should be sensitive to such possibilities.

3.7 Conclusion

This paper accomplishes three tasks. First, it uses a questioning method that tries to minimize social desirability bias in eliciting expressed preferences over transitional justice options. The results were as expected: expressed preferences were less aggressively in favor of punishment and truth-seeking than has been found in a number of recent surveys. Of course, this does not provide definitive evidence of bias in past surveys. Doing so would require randomly assigning different questioning methods in the context of one survey, and doing this on numerous populations. This ought to be the subject of further work for transitional justice survey researchers. Nonetheless, the results from this study raise the question of whether the apparently ardent support for transitional justice interventions suggested by recent surveys are an artifact of the questioning method.

Second, I tested propositions about whether differences in people's expressed preferences for transitional justice options are motivated by differences in political tendencies, levels of insecurity, or levels of knowledge. I find some degree of support for each of these propositions, although the political tendencies hypothesis is most strongly corroborated, particularly with reference to the role of ethnic identity in Burundi's post-war politics. The knowledge proposition finds support only with reference to truth-seeking.

Third, I used an experiment embedded in the survey to study whether expressed preferences were responsive to a simple attempt persuasion. For the most part, it seems not. There was little switching of expressed preferences, and the only clear effects was to make *more resolved* those who take more extreme positions—whether in favor unconditional pardon or unconditional punishment.

These findings suggest that transitional justice interventions ought to be pursued with considerable caution. Gauging support for transitional justice interventions should use methods that aim to elicit preferences with minimal bias. Such has not been the case thus far, in my opinion. Those advocating for transitional justice processes should be deeply aware of how support for such interventions may be based on political calculations as much as principled justice considerations. On the basis of this study, the influence of political calculations appears to apply to both victims and non-victims alike. Given the evidence for the political tendencies hypothesis, those assessing transitional justice needs should also proceed with caution in collecting information on the human rights violations. If the collection of such data is known to anticipate the establishment of punitive mechanisms, then the types of political calculations considered here may affect who comes forward to register abuses. Finally, attempts at persuasion have the potential to backfire, making people more resolute rather than more willing to entertain alternatives. More work should be done on this phenomenon to ensure that deliberation and sensitization interventions do not contribute to polarization.

References

AIHRC Afghan Independent Human Rights Commission (2005) *A Call for Justice: A National Consultation on Past Human Rights Violations in Afghanistan.* Kabul: AIHRC.

Bandura, Albert (1977) "Self-efficacy: Toward a unifying theory of behavioral change." *Psychological Review.* 84(2):191-215.

Bass, Gary J (2001) *Stay the Hand of Vengeance: The Politics of War Crimes Tribunals.* Princeton: Princeton University Press.

BBCWST & SFCG BBC World Service Trust, Search for Common Ground (2007) *Healing a Divided Society: A Survey of Knowledge and Attitudes toward Transitional Justice in Liberia.* London: BBC World Service Trust.

________________ (2008) *Ready to Talk About the Past: A Survey of Knowledge and Attitudes toward Transitional Justice in Burundi.* London: BBC World Service Trust.

Converse, Jean M. & Stanley Presser (1986) *Survey Questions: Handcrafting the Standardized Questionnaire.* Thousand Oaks: Sage.

Freeman, Mark (2010) *Necessary Evils: Amnesties and the Search for Justice.* Cambridge: Cambridge University Press.

Gelman, Andrew (2007) "Struggles with survey weighting and regression modeling (with discussion)." *Statistical Science.* 22(2):153-164.

Gettleman, Jeffrey (2009) "Clinton supports President of Liberia." *New York Times*, 13 August 2009, A9.

Green, Donald P.; Shang E. Ha & John G. Bullock (2010) "Enough already about 'black box' experiments: Studying mediation is more difficult than most scholars suppose." *Annals of the American Academy of Political and Social Sciences*. 628:200-208.

Horowitz, Donald L (1985) *Ethnic Groups in Conflict*. Berkeley: University of California Press.

International Crisis Group (2000) *Burundi: Neither War nor Peace (Africa Report Number 25)*. Brussels: International Crisis Group.

Jackman, Simon & Paul M. Sniderman (2006) "The limits of deliberative discussion: A model of everyday political arguments." *The Journal of Politics*. 68:272-283.

Kaminski, Marek M,; Monika Nalepa & Barry O'Neill (2006) "Normative and strategic aspects of transitional justice." *Journal of Conflict Resolution*. 50(3):295-302.

King, Gary; Robert O. Keohane & Sidney Verba (1994) *Designing Social Inquiry*. Princeton: Princeton University Press.

King, Gary & Langche Zeng (2006) "The dangers of extreme counter-factuals." *Political Analysis*. 14(2):131-159.

Kuran, Timur (1998) *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.

Lemarchand, Rene (1970) *Rwanda and Burundi*. London: Pall Mall Press.

Lie, Tove G.; Helga M. Binningsbo & Scott Gates (2007) "Post-conflict justice and sustainable peace." Typescript, Centre for the Study of Civil War, Peace Research Institute of Oslo, and Norwegian University of Science and Technology.

Lumley, Thomas (2010) *Complex Surveys: A Guide to Analysis Using R*. New York: Wiley.

Morgan, Stephen L. & Christopher Winship. (2007) *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.

Newman, Edward; Oliver P. Richmond & Roland Paris. Eds (2009) *New Perspectives on Liberal Peacebuilding*. Tokyo: United Nations Press.

Noelle-Neumann, Elisabeth (1984) *The Spiral of Silence. A Theory of Public Opinion*. Chicago: University of Chicago Press.

Pankhurst, Donna (1999) "Issues of justice and reconciliation in complex political emergencies: Conceptualising reconciliation, justice and peace." *Third World Quarterly*. 20:239-256.

Pham, Phuong N.; Harvey M. Weinstein & Timothy Longman (2004) "Trauma and PTSD symptoms in Rwanda: Implications for attitudes toward justice and reconciliation." *Journal of the American Medical Association*. 292(5):602-612.

Pham, Phuong N.; et al (2005) *Forgotten Voices: A Population-Based Survey on Attitudes About Peace and Justice in Northern Uganda*. Berkeley: Human Rights Center.

Pham, Phuong N.; et al (2007) *When the War Ends: A Population-Based Survey on Attitudes About Peace, Justice, and Social Reconstruction in Northern Uganda*. Berkeley: Human Rights Center.

Rotberg, Robert I. & Dennis Thompson. Eds (2000) *Truth v. Justice*. Princeton University Press.

Samii, Cyrus (Nd) "*Wartime and Post-conflict Experiences in Burundi* technical note: Hierarchical Bayes small area estimates of commune level ethnicity proportions." Typescript, Columbia University. Available at http://www.columbia.edu/∼cds81/l

Snyder, Jack L. & Leslie Vinjamuri (2003) "Trials and errors: Principles and pragmatism in strategies of international justice." *International Security.* 28(3):5-44.

Sudharshan, Ramaswamy (2003) "Rule of law and access to justice: Perspectives from UNDP Experience." Unpublished paper presented to the European Commission Expert Seminar on Rule of Law and Administration of Justice as part of Good Governance, Brussels.

Theidon, Kimberly (2006) "Justice in transition: The micropolitics of reconciliation in postwar Peru." *Journal of Conflict Resolution.* 50(3):433-457.

Thoms, Oskar N.T.; James Ron & Roland Paris (2008) "The effects of transitional justice mechanisms. A summary of empirical research findings and implications for analysts and practitioners." Typescript. Center for International Policy Studies, University of Ottawa.

United Nations (1996) *Final Report of the International Committee of Inquiry for Burundi.* New York: United Nations.

United Nations (2004) *The Rule of Law and Transitional Justice in Conflict and Post-Conflict Societies. Report of the Secretary General, S/2004/616.* New York: United Nations.

UNDP United Nations Development Programme (2007) *Public Perceptions on Transitional Justice: Report on Transitional Justice Opinion Polling Survey Conducted in April-May 2007 in Kosovo.* Pristina: UNDP Kosovo.

United Nations Office of the High Commissioner for Human Rights (2006) *Rule of Law Tools for Post-Conflict States, Documents HR/PUB/06/1 - HR/PUB/09/2.* New York and Geneva: United Nations.

Vinck, Patrick; et al (2008) *Living with Fear: A Population-Based Survey on Attitudes About Peace, Justice, and Social Reconstruction in Eastern Democratic*

Republic of Congo. Berkeley: Human Rights Center.

Wimmer, Andreas (2006) "Ethnic exclusion in nationalizing states." In Gerard Delanty & Krishan Kumar, eds. *Handbook of Nations and Nationalism.* London: Sage.

Tables

Table 3.1: **Summary statistics**

Variable	Mean	Wgtd. Mean[a]	S.D.	Possible values	N
Punishment:				1,2,3	1162
i. Uncond forgive	0.36[b]	0.34[b]			
ii. Cond forgive	0.56[b]	0.60[b]			
iii. Uncond punish	0.07[b]	0.05[b]			
Seek truth	0.39	0.31	0.49	0,1	1155
Non-Hutu	0.29	0.26	0.45	0,1	1169
Southern	0.19	0.15	0.39	0,1	1169
Non-Hutu X Southern	0.05	0.0	0.22	0,1	1169
Higher ed.	0.31	0.29	0.46	0,1	1169
Victim. Reb.	0.22	0.26	0.42	0,1	1169
Victim. FAB	0.24	0.23	0.43	0,1	1169
Insecurity	0.28	0.26	0.45	0,1	1169
InsecurityXVictim. Reb.	0.06	0.07	0.24	0,1	1169
InsecurityXVictim. FAB	0.09	0.09	0.29	0,1	1169

[a]Mean with survey weights applied.

[b]Proportions for the punishment preference variable.

Table 3.2: **Mass surveys on transitional justice attitudes**

Authors	Pub. year	Location	Pop.	Research year	Sample size	Sampling method	Percent stating justice top priority	Modal responses to questions about TJ processes		
								General accontability	Punishment	Truth
AIHRC	2005	Afghanistan	Adults nationwide	2004	4,151	Enumerator discretion in designated locations with preset criteria	4% say rule of law is top priority.	49%: justice means legal punishment. 76%: justice very important.	76%: punishment of war criminals contributes to security. 28%: only serious offenders and commanders should be tried. 61%: reject amnesty of war criminals.	95%: important to establish record of truth. 23%: Govt should take lead in establishing record.
BBCWS T & SFCG	2007	Liberia	Adults in 8 of of 15 counties	2007	1,600	Villages/towns selected to reflect diversity, particularly ethnic diversity. Households selected by enumerators quasi-randomly within villages/towns. Stratification by gender.	29% mention access to courts as one among a number of priorities.		92%: recall human rights abuses during wartime. 50%: resondents who could recall abuses thought that abusers should go to trial. 60%: commanders or faction leaders should be tried. 42%: of those aware of TRC thought it should use amnesty power. 44%: of those aware of TRC thought it should not use prosecutorial power.	79%: aware of TRC. 56%: of those aware of TRC think its contribution to reconciliation good or excellent. 46%: of those aware of TRC think its contribution to truth has been good or excellent. 36%: of those aware of TRC think its contribution to accurate account has been good or excellent.
BBCWS T & SFCG	2008	Burundi	Adults in 10 out of 17provinces	2008	1,648	Communities selected to reflect diversity. Households selected by enumerator quasi-randomly within communities. Stratified by age group.	19% mention trying human rights offenders and 11% mention establishing truth as one among a number of priorities.	36%: Justice means fair treatment.	68%: wrongdoers should be brought to trial. 48%: aware of proposals for special court.	68%: aware of proposal for TRC. 81%: think TRC wold be entirely or mostly good. 59%: of those thinking TRC good say it's because of reconciliation.
Pham et al	2004	Rwanda	Adults in 4 of 154 communes	2002	2,091	Communities selected to reflect diversity. Respondents selected randomly within communes.			42%: support international tribunal. 68%:support domestic trials. 91%: support gacaca.	
Pham et al	2005	Northern Uganda	Adults in 4 highly war-exposed districts.	2005	2,585	Districts selected based on high conflict exposure, and ethnic and language diversity. IDP camps and other sites stratified by urbanness selected in districts. Respondents selected randomly within sites/camps.	Less that 1% saying justice is a concern over other development/security priorities.	31%: justice means trials. 76%: LRA human rights abusers should be held accountable. 76%: UPDF human rights abusers should be held accountable. 58%: lower level LRA offenders should not be held accountable.	66%: LRA rights offenders should be tried and punished. 51%: UPDF rights offenders should be tried and punished. 53-80%: rights offenders should have amnesty. 65%: accept amnestied LRA leaders if they returned home. 79%: sccept amnestied low level LRA if they returned home. 56%: Require amnestied LRA to apologize for forgiveness 56% vs. 39%: peace with amnesty preferred to peace with trials, in Acholi vs. non-Acholi districts.	92%: truth telling process needed in Northern Uganda.
Pham et al	2007	Northern Uganda	Adults in 8 highly war-exposed districts.	2007	2,875	Districts selected based on high conflict exposure, and ethnic and language diversity. IDP camps and other sites stratified by urbanness selected in districts. Respondents selected randomly within sites/camps.	3% say justice is top priority.	41%: justice is being fair 33%: prevent future comflict by pardoning LRA leaders 70%: important to hold human rights offenders accountable 50%: hold LRA leaders accountable 48%: hold all LRA accountable 40%: hold government accountable 18%: hold UPDF accountable.	32%: LRA rights offenders should be tried and punished. 55%: UPDF rights offenders should be tried and punished. 59%:agree important to have trials for LRA leaders 34%:important to have trials for lower rankinng LRA 78%: lower level LRA should be pardoned 80%: peace with amnesty preferred to peace with trials	95%: establish written record. 90%: establish truth commission.
UNDP	2007	Kosovo	Adults in	2007	1,250	Stratification on demographic, geographic, and ethnic features. Individual selection method unclear.			36-70%: satisfaction with ICTY for K-Serbs vs. Kalbanians 90%: punishment of perpetrators crucial element of justice	86-83%: believe establishing trust is important for K-Albanians vs K-Serbs
Vinck et al	2008	Eastern DRC	Adults in areas most affected by war: Ituri, North Kivu, and South Kivu districts.	2007	2,620	Communities and individuals randomly selected.	3% say justice is top priority.	85%: important to hold offenders accountable. 82%: acountability necessary to secure peace. 51%: justice means establishing truth.	69%: war criminals should be punished. 38%: foot soldiers should be treated the same as leaders.	88%: important to know the truth about what happened. 56% - 24%: establish truth via judiciary vs. truth commission.

Table 3.3: **Demographic characteristics of sample**

Demographic category		Raw sample %	Weighted sample %
Gender:			
	Men	86%	47%
	Women	14%	53%
Ethnicity			
	Hutu	71%	74%
	Tutsi	28%	25%
	Other	1%	1%
Highest level of education			
	Primary not completed	43%	46%
	Primary	37%	35%
	Junior secondary	11%	11%
	Senior secondary	7%	7%
	University+	2%	1%
Sample total: 1,169 civilians			

Note: The male-female ratio was by design. Refer to the text.

Table 3.4: **Overall preference distribution**

	Unconditional forgiveness	Conditional forgiveness	Unconditional punishment	Total
Seek truth about past	7% (1)	22% (2)	2% (1)	31% (3)
Forget the past	28% (3)	39% (3)	3% (1)	69% (3)
Total	34% (3)	60% (3)	5% (1)	

Cell proportions displayed.
N = 1151
Percentage point standard errors are shown in parentheses.
Rao-Scott F, 3.97; p-value, 0.03.

Note: The data contained missing values for seven respondents on the punishment question and 14 respondents on the truth question. These were simply omitted.

Table 3.5: **Ordered logistic regression of preference for punishment, odds ratio estimates**

Variable		**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Non-Hutu	*OR*	1.61	2.00	1.95	1.69	1.60
	(s.e.)	(0.41)	(0.55)	(0.53)	(0.45)	(0.43)
	p-val	0.07	0.02	0.02	0.06	0.09
Southern province	*OR*		1.01	1.02	0.94	0.91
	(s.e.)		(0.29)	(0.29)	(0.25)	(0.25)
	p-val		0.95	0.95	0.81	0.74
Non-Hutu X South.	*OR*		0.26	0.25	0.3	0.31
	(s.e.)		(0.20)	(0.19)	(0.22)	(0.23)
	p-val		0.09	0.08	0.12	0.13
Higher education	*OR*			1.09	1.05	1.09
	(s.e.)			(0.31)	(0.29)	(0.31)
	p-val			0.76	0.86	0.78
Rebel victimization	*OR*				1.89	2.27
	(s.e.)				(0.49)	(0.67)
	p-val				0.02	0.01
Army victimization	*OR*				1.04	0.9
	(s.e.)				(0.27)	(0.31)
	p-val				0.88	0.76
Insecurity	*OR*					0.92
	(s.e.)					(0.34)
	p-val					0.83
Insecurity X Rebel vict.	*OR*					0.55
	(s.e.)					(0.34)
	p-val					0.34
Insecurity X Army vict.	*OR*					1.59
	(s.e.)					(0.89)
	p-val					0.41
N		1162	1162	1162	1162	1162
Global F-test p-val.		0.07	0.03	0.07	0.04	0.17

Note: "OR" refers to "odds ratio." Standard errors are on the odds ratio scale, accounting for stratification at the province level and clustering at the commune level. The data contained missing values for seven respondents. These were omitted.

Table 3.6: **Logistic regression of preference for truth-seeking, odds ratio estimates**

Variable		**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Non-Hutu	*OR*	1.86	1.87	1.53	2.10	1.94
	(s.e.)	(.47)	(.54)	(.45)	(.64)	(0.56)
	p-val	0.02	0.04	0.15	0.02	0.03
Southern	*OR*		1.59	1.66	1.95	1.93
province	*(s.e.)*		(0.58)	(.63)	(.71)	(0.74)
	p-val		0.20	0.19	0.07	0.09
Non-Hutu	*OR*		1.03	0.88	0.68	0.70
X	*(s.e.)*		(.53)	(.46)	(.32)	(0.35)
South.	*p-val*		0.95	0.81	0.42	0.47
Higher	*OR*			2.33	2.47	2.55
education	*(s.e.)*			(.60)	(.63)	(0.63)
	p-val			0.00	0.00	0.00
Rebel	*OR*				0.46	0.61
victimization	*(s.e.)*				(.13)	(0.21)
	p-val				0.01	0.16
Army	*OR*				1.37	1.20
victimization	*(s.e.)*				(.43)	(0.34)
	p-val				0.33	0.52
Insecurity	*OR*					1.98
	(s.e.)					(0.67)
	p-val					0.05
Insecurity	*OR*					0.35
X	*(s.e.)*					(0.19)
Rebel vict.	*p-val*					0.06
Insecurity	*OR*					1.11
X	*(s.e.)*					(0.71)
Army vict.	*p-val*					0.87
N		1155	1155	1155	1155	1155
Global F-test p-val.		0.02	0.01	0.00	0.00	0.00

Note: "OR" refers to "odds ratio." Standard errors are on the odds ratio scale, accounting for stratification at the province level and clustering at the commune level. The data contained missing values for seven respondents. These were simply omitted.

Table 3.7: **Persuasion experiment results**

Initial response: Unconditional forgiveness			
	Unconditional forgiveness	Conditional forgiveness	Unconditional punishment
Vacant counter-argument	76.7%	18.7%	4.6%
	(4.6)	(3.9)	(4.4)
Content-laden counter-argument	83.1%	16.8%	0.1%
	(4.0)	(4.0)	(0.1)
Total	79.9%	17.8%	2.4%
	(3.3)	(3.0)	(2.2)

Row proportions displayed.
Subgroup N = 404
Percentage point standard errors are shown in parentheses.
Rao-Scott F, 4.28; p-value, 0.03

Initial response: Conditional forgiveness			
	Unconditional forgiveness	Conditional forgiveness	Unconditional punishment
Vacant counter-argument	6.5%	88.1%	5.4%
	(2.9)	(3.7)	(2.6)
Content-laden counter-argument	9.4%	84.1%	6.5%
	(2.9)	(3.8)	(2.8)
Total	7.9%	86.1%	6.0%
	(2.1)	(3.0)	(2.4)

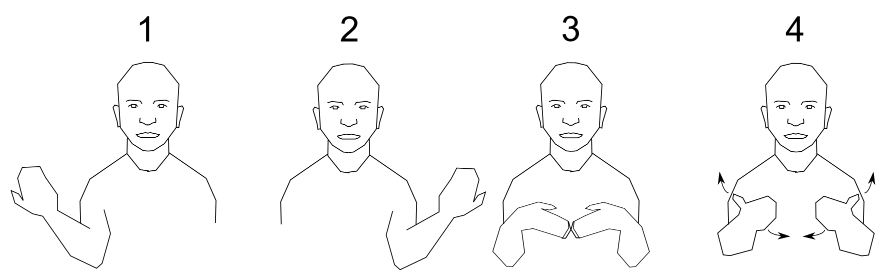
Row proportions displayed.
Subgroup N = 664
Percentage point standard errors are shown in parentheses.
Rao-Scott F, 0.40; p-value, 0.65

Initial response: Unconditional punishment			
	Unconditional forgiveness	Conditional forgiveness	Unconditional punishment
Vacant counter-argument	0.3%	24.5%	75.2%
	(0.3)	(11.7)	(11.7)
Content-laden counter-argument	0.0%	2.4%	97.7%
	(0.0)	(1.7)	(1.7)
Total	0.1%	12.2%	87.7%
	(0.1)	(5.8)	(5.9)

Row proportions displayed.
Subgroup N = 78
Percentage point standard errors are shown in parentheses.
Rao-Scott F, 10.20; p-value, 0.00

Figures

Figure 3.1: **Gestures accompanying question delivery**



Punishment question:

Gesture 1. Some people say that former combatants who killed civilians or raped women should not be accepted in their home communities in any case and they should be punished [choice 1].

Gesture 2. Some other people say that they should be accepted and what happened should be forgotten [choice 2].

Gesture 3. A third group says that they could be accepted if they beg for forgiveness [choice 3].

Gesture 4. Which one of the three groups do you support?

Truth-seeking question:

Of the following options, which is close to your point of view?

Gesture 1: In order to achieve peace and reconciliation, it is necessary to know the truth about what happened before the war [choice 1].

Gesture 2: In order to achieve peace and reconciliation, it is good to forget about the past [choice 2].

Figure 3.2: **Persuasion experiment design and treatment group sizes**

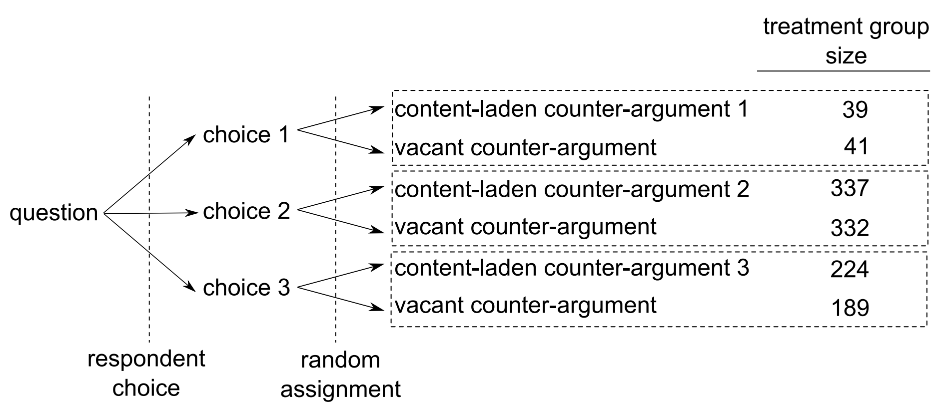
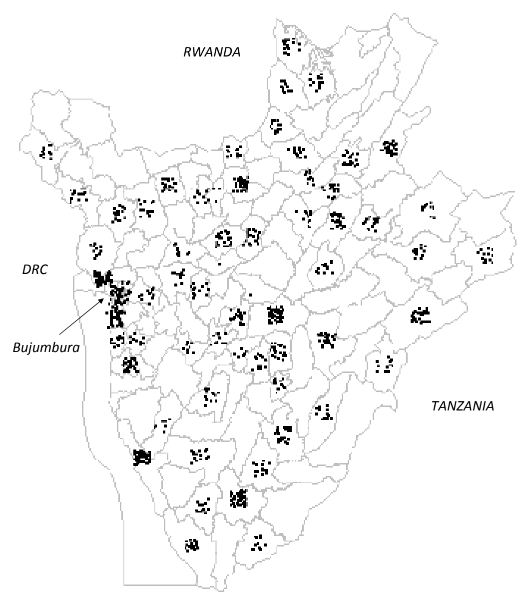


Figure 3.3: **Geographic locations of survey respondents**



Chapter 4

Reintegrating rebels into civilian life: Quasi-experimental evidence from Burundi

Summary

We use original survey data, collected in Burundi in the summer of 2007, to show that a World Bank ex-combatant reintegration program implemented after Burundi's civil war caused significant economic reintegration for its beneficiaries but that this economic reintegration did not translate into greater political and social reintegration. Previous studies of reintegration programs have found them to be ineffective, but these studies have suffered from selection bias: only ex-combatants who self selected into those programs were studied. We avoid such bias with a quasi-experimental research design made possible by an exogenous bureaucratic failure in the implementation of program. One of the World Bank's implementing partners delayed implementation by almost a year due to an unforeseen contract dispute. As a result, roughly a third of ex-combatants had their program benefits withheld for reasons unrelated to their reintegration prospects. We conducted our survey during this period, constructing a control group from those unfortunate ex-combatants whose benefits were withheld.

We find that the program provided a significant income boost, resulting in a 20 to 35 percentage point reduction in poverty incidence among ex-combatants. We also find moderate improvement in ex-combatants' livelihood prospects. However, these economic effects do not seem to have caused greater political integration. While we find a modest increase in the propensity to report that civilian life is preferable to combatant life, we find no evidence that the program contributed to more satisfaction with the peace process or a more positive disposition toward current government institutions. Reintegration programs are central in current peace processes and considerable resources are devoted to them. Thus, our evidence has important policy implications. While we find strong evidence for the effectiveness in terms of economic reintegration, our results challenge theories stating that short-run economic conditions are a major determinant of ones disposition toward society and the state. Social and political integration of ex-combatants likely requires much more than individually-targeted economic assistance.

4.1 Introduction

Are post-conflict disarmament, demobilization, and reintegration programs effective in achieving their programmatic aims of economic rehabilitation? Does such economic assistance make ex-combatant beneficiaries more likely to reintegrate politically and socially? We address these questions, exploiting quasi-experimental evidence from an ex-combatant reintegration program implemented in Burundi after the 1993-2004 civil war.

These questions are important for policy and for the study of transitions from war to peace. Post-conflict DDR programs have become key features of most peace agreements that end civil wars.[1] According to the 2006 Uppsala Peace Agreement dataset, 36% of peace agreements struck in 1989-1999 period contained DDR provisions, while in 2000-2005, the percentage rose to 59%.[2] DDR programs are typically carried out alongside economic reconstruction, refugee repatriation, and democratization. All of these "peacebuilding" interventions aim to address the root causes of conflict and reduce the likelihood of another war (Boutros-Ghali, 1992; Cousens et al, 2001; Paris, 2004; Doyle and Sambanis, 2006). DDR programs are usually implemented under the direction of international actors, especially the United Nations (UN) and the World Bank. DDR programs encompass a set of activities designed to lay the foundation of stable and self-sustaining peace. These activities entail the discharge of a large number of combatants who must be disarmed and reintegrated into civilian life (Walter 2001; Fortna 2008).

---

[1]Not all DDR take place in the framework of peace agreements. In several cases, like Rwanda and Ethiopia, DDR was undertaken after the military defeat of one side. In Colombia, DDR was launched while the civil war was ongoing. Some DDR programs take place as part of army reform processes, such as those that followed World War II and the Cold War (Kingma and Pauwels 2000). This paper focuses on internationally-sponsored DDR in post-conflict settings.

[2]The Uppsala Peace Agreement dataset is available online at http://www.pcr.uu.se/research/UCDP/.

The disarmament and demobilization of former combatants—the first two components of DDR—take place before the reintegration phase in order to create security conditions and trust necessary for implementing peace agreements (Spear 2002). Disarmament and demobilization are a prerequisite for the roll-out of recovery and development activities (Kumar 1997; Rubin et al. 2003; Feil 2004). A goal following disarmament and demobilization is for ex-combatants to find a livelihood and submit to laws and norms that govern civilian society. The fear is that demobilized combatants may have difficulty in finding a productive position in the legal civilian economy and may maintain an oppositional stance toward society and government. Such marginalization may increase the propensity to engage in crime or even renewed organized violence. As Alden (2002) remarks, "the spectre of former military personnel in criminal networks in the Balkans and Russia, the outbreak of violence inspired by real and self-proclaimed war veterans in Zimbabwe, and the participation of former security force members from Eastern Europe and South Africa in mercenaries in war-torn Angola and the Congo serve to underscore the destabilizing role played by former combatants who remain outside of the economy and society as a whole." Spear (2006) argues that among post-conflict interventions, reintegration is the most directly "linked to establishing a *lasting* peace" (emphasis in the original)—a conclusion echoed by Cavallo's (2008) argument that effective reintegration of former combatants is a "*sine qua non*" for the consolidation of peace.

By providing a range of economic benefits, reintegration programs try to make civilian life more attractive to ex-combatants and thus reduce the risk of political disorder. The presumption is that the former combatants cannot or will not achieve productive civilian livelihoods *automatically*. This might be due to economic barriers, such as lack of access to productive resources or a soft labor market, as well as social factors, such as an oppositional disposition from the ex-combatants themselves or discrimination by others. Reintegration programs usually include (i) short-term measures such as the provision of cash assistance or in-kind material benefits to address the immediate needs of former combatants and their dependents upon leaving armed groups; and (ii) longer-term measures such as vocational training, skill development and counseling, which aim to reintegrate former combatants into the social and economic structures of society (Colletta et al. 1996; Bryden et al.2005; Muggah 2009). The literature that informs the design of current reintegration programs tends to presume that ex-rebels form a class that, on average, exhibits material, human, and social capital deficiencies. If these deficiencies are not attended to, the presumption is that they would significantly limit ex-combatants' ability to establish themselves economically. making them susceptible to recruitment into armed groups or criminality. That is, with few opportunities to establish a productive livelihood in the legal civilian sector, ex-combatants will face difficulties in releasing their conflict-era identities and may find it difficult to resist the lure of using coercion and promotion of instability to satisfy their material and psychological needs. Sometimes program designers may appreciate that ex-combatants are no more vulnerable than their civilian counterparts, but that they need extra incentives to be steered from the temptations of banditry and extortion—activities that their combatant experience has trained them to do. Finally, in some contexts (e.g. where rebels have won control of the government), reintegration assistance is justified as necessary for honoring the sacrifices of combatants who were not integrated into a new armed forces.

We continue below by discussing findings from related literature. We then discuss in more detail outcomes of interest and hypotheses pertaining to reintegration programs. We describe the context of rebel reintegration in Burundi after the 1993-2004 civil war, and explain the shock to program access that we exploit to measure program effects. We describe the methods that we use to identify effects of the program—or more correctly, effects of the *non-availability* of the program—and then present our estimates of effects on economic and political reintegration. We discuss these results and then conclude with suggestions for further research.

4.2 Related literature

Despite the importance of reintegration programs, there have been few attempts to measure their effectiveness. Some macro-level studies exist, particularly studies evaluating the reintegration programming of major international institutions like United Nations agencies and the World Bank, and they typically focus on program performance and outputs.[3] There are several descriptive studies that focus on the practical challenges that fraught the implementation of reintegration programs,[4] a growing critical literature that questions the assumptions that underpin these programs,[5] as well as numerous qualitative studies that offer detailed accounts of beneficiaries' perspectives.[6] While useful, these studies do not measure the benefits of reintegration programs in a manner necessary for assessing their gross value relative to the costs. These costs include diverting resources from others in need. In addition, reintegration programs are often understood as unjustifiably "rewarding" former combatants for engaging in violence.[7]

Two recent studies have tried to measure the benefits of reintegration programs, using large scale surveys of ex-combatants. Humphreys and Weinstein (2007) investigate the effects of a UN-sponsored DDR program in Sierra Leone. They used propensity score matching on ex-combatants who did and did not *enroll* or *complete* the program. For both types of participation, these comparisons yielded effect estimates centered on zero for economic reintegration and political reintegration. These authors qualify their results by stating that "no evidence of an effect" is not

---

[3]See, for example, Colletta et al (1996) and the studies contributed to the Government of Sweden's *Stockholm Initiative on Disarmament, Demobilization, and Reintegration* (2006).

[4] See Kingma (1997); UNDP, 2000; Alden (2002); Paes (2005); Baare (2006) and Specker (2008).

[5]See Pouligny (2004); Muggah (2005) and Jennings (2008).

[6]See Jennings (2007), Soderstrom (2010), and Uvin (2007).

[7]The introduction to Muggah (2009) summarizes the on-going debate among policy-makers and academics about the utility of reintegration programs as currently designed.

necessarily "evidence of no effect." Nonetheless, the findings from this study have raised questions about whether reintegration programs accomplish *anything*. Pugel (2009) conducted a similar study among ex-combatants in Liberia. Using a regression analysis to control for background characteristics, Pugel found that those who had *completed* the UN reintegration program were significantly more likely to have a "livelihood-producing activity," although no statistically significant effects were evident for ex-combatants' likelihood of earning an income above the poverty line or "exhibiting spending patterns indicative of excess earning capacity."[8] Pugel found no significant effects associated with other forms of participation in the program. In a re-analysis of Pugel's data using adjustment techniques similar to those of Humphreys and Weinstein (2007), Levely (2010) found no effects of *registration* into the reintegration program on income (confidence intervals were centered near zero), but he found some indication that *enrollment* and *completion* of the program increased income (on the order of a 10% increase), although among the many estimates he computed, most of the 95% confidence intervals overlapped with zero. These findings raise questions about whether reintegration programs, as currently designed, are effective in achieving the immediate goal of economic reintegration, much less the downstream goal of political reintegration.

As the authors themselves admit, these studies cannot be definitive on the effectiveness of reintegration programs in Liberia and Sierra Leone, much less in general, because of the weak possibilities that they offer for identifying causal effects. The weakness is by no fault of the authors, but rather by the nature of the programs. The material benefits offered by such programs makes the incentive to participate very strong, meaning that those who do not participate are likely to differ in important ways from those who do. In Sierra Leone, for example, rates of enrollment in the UN-sponsored reintegration program at the time of Humphreys and Weinstein's fieldwork were around 90 percent (Humphreys and Weinstein, 2007), in which case the

[8]Pugel (2009) only presents the signs and statistical significance of his estimates.

non-participants used to construct a pseudo "control group" are likely to be a highly self-selected group. In Liberia, the rates of enrollment at the time of Pugel's field-work were lower—about 25%—but Pugel himself indicates that such enrollment was voluntary. If self-selection is the main reason for variation in treatment assignment, then even among subjects matched on observed covariates, there is strong reason to remain concerned about bias due to unobserved prognostic factors associated with program take-up.[9]

Sekhon (2009) argues persuasively that adjustment methods, such as matching or regression, should be tied *nonetheless* to a discontinuity or exogenous shock to make the causal story persuasive. A convincing causal interpretation of an adjusted estimated requires that we can explain how two individuals who share similar observable traits may differ in their program status for reasons that will not bias the analysis. The clearer we can be about the source of such exogenous variation, the more convincing the interpretation. In the absence of some random shock to program access, identifying impacts of reintegration programs is a task that requires more assumptions than most people would be comfortable to make. In this paper we exploit such a shock, which occurred during the reintegration program in Burundi. This allows us to measure reintegration program effects with minimal taint from self-selection bias.

4.3 General hypotheses about reintegration programs

The United Nations defines reintegration as "the process which allows ex-combatants and their families to adapt, economically and socially, to productive civilian life" (United Nations, 2000). Reintegration is critical in peace processes because it links "the more immediate requirements of disarmament and demobilization to the long-

[9] DiNardo and Lee (2010) discuss conditions under which matching can even exacerbate self-selection bias.

term imperatives of social and economic welfare" (Bryden et al.2005) and because it is the set of activities that facilitate effective conversion from combatant to civilian life. Effective conversion to civilian life is achieved not only when former combatants are able to establish peaceful and sustainable livelihoods ("economic reintegration"), but also when they longer see violence as a legitimate means to seek political or personal gain and they submit to the laws and norms of civilian society ("political reintegration"; World Bank, 2002; United Nations, 2000; Pouligny 2004; IDDRS 2010).[10] The ability to establish good working relations with one's community or family ("social reintegration") is often taken to be an important moderator of an ex-combatant's economic and political reintegration prospects. The typical causal model posited in reintegration program documentation is economic reintegration → political reintegration, with family- and community-level social reintegration moderating the process. Reintegration programs typically emphasize economic reintegration as their main *programmatic* objective, but political reintegration is usually the ultimate goal. Reintegration programs may also include reconciliation interventions that attempt to facilitate social reintegration, but these are usually secondary. Political reintegration is not directly targeted but presumed to follow from the boost that a program provides to social and economic reintegration. In this way, the programmatic economic reintegration effects are presumed to have down-stream political reintegration effects.[11]

The emphasis on enhancing ex-combatants' economic welfare draws in part on the research program on the political economy of conflict (Collier and Hoffler 2004; World

[10] Rackley (2007) argues that donors have the mistaken idea that "as soon as you get guns out of their hands, they are suddenly innocuous human beings again, but that is not the case at all" (cited in Hanson 2007). To complete the conversion process, as Rackley discusses, former combatants must be provided with an economic alternative to living by the gun.

[11] This implicit "theory" of reintegration is well articulated in World Bank (2004a), which layed the foundation for recent reintegration programming in Angola, Burundi, Central African Republic, Democratic Republic of Congo, Republic of Congo, Rwanda, and Uganda.

Bank, 2002; 2004a; Spear, 2006). In its 2003 report entitled *Breaking the Conflict Trap*, the largest sponsor of reintegration programs, the World Bank, claimed that "a structured DDR process, which demobilizes combatants in stages and *emphasizes their ability to reintegrate into society*, may reduce the risk of ex-combatants turning to violent crime or joining rebel groups in order to survive" (emphasis added). Ex-combatants may lack human and social capital endowments necessary to establish themselves in the peaceful economic sector.[12] These deficiencies may result from characteristics that set apart those who decided to participate in fighting, or they may be consequences of soldiering itself. Ex-combatants may face acute shortages of material capital. Time spent away soldiering may have caused an ex-combatant to have had his or her land taken over by someone else; or, time spent soldiering may have made it difficult for the person to acquire land in the first place. These deficiencies may put ex-combatants at an economic disadvantage and make them more likely to shirk the law in finding ways to support themselves. This suggestion is consistent with findings from Collier's (1994) research on demobilization and reintegration of former soldiers in Uganda.[13] Some combatants may have joined armed groups because the economic opportunity costs associated with doing so were low. Thus, as Spear (2006) notes "an emphasis on [economic reintegration] recognizes that some of the motives for fighting were economic and that if the economic dimensions of the problem are

[12]For broad-brush reviews on possible sources of vulnerability among ex-combatants, see Annan and Patel (2009) and Tajima (2009).

[13]Some analysts suggest that the problem of postwar reintegration is not so much that ex-combatants have human and social capital deficiencies, but rather that the local economies they reintegrate *into* have very few economic opportunities available for them. For instance, speaking of the economic reintegration challenges faced by ex-combatants in Mozambique, McMullin (2004) notes "post-conflict states with impoverished economies offer little to reintegrate into. Mozambique, where only a tenth of the population had formal employment, was not exception, giving rise to quip: 'the government told us "now you are all equally poor. You have been reintegrated back into basic poverty" (emphasis in original).

not addressed, any settlement of the conflict may be short-lived." The absence of law and order in conflict situations may provide combatants with opportunities to enrich themselves through illicit means, including robbery, racketeering and smuggling (Reno 1997). Ending the war is tantamount to giving up these opportunities. Anecdotal evidence suggests that some combatants make these calculations. During the DDR process in Liberia, a member of the armed group, Liberians United for Reconciliation and Democracy (LURD) told a reporter, "I still have my 81-mm mortar, but I just come to see whether the UN was giving fighters who disarm something good, If they don't give good money, I will not give the rocket" (Agence France Press, 8 December 2003; cited in Spear 2006). Finally, grievances over lack of access to opportunities may have motivated a willingness to challenge the state. Providing opportunities in the legal economy may remove this motivation and, as a result, an anti-state orientation.

In our empirical analysis we test two interrelated hypotheses:

Programmatic hypothesis The first hypothesis is that *reintegration programs substantially improve the economic welfare of ex-combatants*, as they are programmatically designed to do.

Downstream hypothesis The second hypothesis is that *by improving economic welfare, reintegration programs substantially increase ex-combatants' willingness to respect the rule of law and adopt an orientation that favors societal stability.*

This paper investigates these hypotheses with quasi-experimental evidence in Burundi. We look at the extent to which the reintegration program achieved its proximate aims of improving income and livelihood outcomes. We then look at whether the programs achieved its downstream aims of increasing ex-combatants' satisfaction with civilian life relative to combatant life, inducing a more positive orientation toward the peace agreement, and inducing a more positive orientation toward stable government institutions.

4.4 The Burundi context

We focus on the reintegration of adult (18 years old or more) male former rebel combatants in Burundi. Members of the national army were also demobilized and offered reintegration assistance, but we do not study them. The reintegration experience of former army members is likely to be quite different than for former rebels. Being an army soldier is a legal, well defined career. Demobilization and reintegration support is well institutionalized: certain members of the national army had access to a system of pension-type benefits that were separate to those put in place by the internationally assisted reintegration program. The legality and institutionalized nature of an army career implies that there are fewer questions about demobilized army soldiers' "place in society." For these reasons, we do not think it is warranted to lump the two subgroups together. We choose to focus on reintegration of former rebels, which we believe to be of utmost interest to reintegration program designers. We focus on men only because we think women's experiences are likely to be distinct, but our sample of women is too small to study them adequately. Finally, we focus on former rebels who were aged 18 years or older as of fieldwork in 2007. Some of them were recruited before adulthood,[14] but they were all adults at the time of their demobilization, which could have been up to 3 years prior to fieldwork. Ex-combatants under 18 were treated by a different (UNICEF-managed) reintegration program than adults, in which case outcomes associated with them would not necessarily be comparable to those above the age cut-off.[15]

[14]The youngest ex-rebel in our sample is 19, and the years of recruitment for our respondents was between 1993 and 2003.

[15]Such a strict age cut-off creates an opportunity for a possibly well-identified examination of the benefits of programming for those just under 18 relative to those just over 18. This is something that could be pursued in further work. Those under 18 received considerably more support relative to their counterparts over 18 in mending relations with family and community members. Thus, there is a ripe opportunity for studying the impacts on social reintegration and the downstream benefits

4.4.1 Background to the reintegration program

The DDR program in Burundi was initiated following a 2003-4 ceasefire that drew into the peace process the largest rebel group in the country, the National Council for the Defense of Democracy-Forces for the Defense of Democracy, or CNDD-FDD by its French acronym. At the time of our fieldwork in 2007, one rather small rebel faction—the Agathon Rwasa faction of the National Forces for Liberation, or FNL by their French acronym—remained outside the peace process. The war had begun in 1993 in aftermath of a tumultuous attempt at democratization. Elections in 1993 had resulted in the triumph of a party that represented the aspirations of a long-oppressed Hutu majority. Under still-mysterious circumstances, members of the southern- and Tutsi-dominated army led a failed coup attempt in October 1993; the coup attempt nonetheless involved the assassination of the recently-elected president. The event triggered massive violence throughout the country. The ensuing ferment gave way to outright civil war. The war had a devastating impact on the tiny, landlocked central-African country. Fighting touched most of the country. It resulted in deaths estimated at approximately 300,000 out of a total population of 6-8 million. Burundi's pre-war socio-economic development levels were already among the world's lowest, although for its income level, the country did have relatively well-developed infrastructure and institutions. The war severely stalled development for over a decade, resulting for example in an estimated 20% decline in real GDP over 1993-2002 (World Bank 2004, p. 6).

The outcome of the war and ensuing political developments are important features of the environment into which ex-rebels were reintegrating. The war resulted in a peace accord between the government and the CNDD-FDD that called for elections. Significantly, the 2005 elections resulted in the CNDD-FDD winning an outright majority of national assembly seats (59% of 118 seats) and communal councilor posts (55% of 3,225 posts). This gave the party the strength necessary to elect its political

---

of social reintegration on economic and political reintegration.

head, Pierre Nkurunziza, to be president. As such, the outcome of the war brought about a near revolution in the institutionalized political context relative to before the war. As of 2005, the former rebels were elected to lead. Contrast this to outcomes in the other countries where reintegration has been studied: in Sierra Leone, the party of former rebels, the Revolutionary United Front Party, managed barely 2% of votes, failing to win a single seat in the 2002 elections that punctuated the end of the war. In Uganda, the Lord's Resistance Army is still an illegal organization, never having had its popular legitimacy tested. Though the number of cases is too small to test the claim rigorously, there is good reason to believe that conditions causing rebel parties to fair well electorally are associated with reintegration prospects of demobilized rebels. These considerations should be taken into account when trying to generalize the findings from this paper.

The 2000 Arusha Accords called for DDR, and these Accords set the parameters of the 2003 Pretoria agreement, which brought the CNDD-FDD into the peace process and set the stage for the 2005 elections. A February 2004 World Bank report set specific DDR targets; a November 2004 Joint Operations Plan (JOP), which was approved by a committee that included all key national and international agencies, made these targets official (World Bank 2004; Republic of Burundi 2004; Boshoff and Vrey 2006, p. 15). The MDRP and national implementing agencies designed a program that was intended to demobilize and reintegrate enough combatants to achieve a new army (Forces de Defense Nationale) of 25,000 and police force of 20,000.[16] The

[16]The original project proposal set a demobilization target of 55,000 combatants. However, this target appears to have been an error: the size of the various forces prior to demobilization was estimated to be about 80,000. The target army plus police force size was 45,000. If no new recruitment was to take place, somewhere around 35,000 would need to be demobilized. Recruitment into the new armed forces was to take place, but no where near the 20,000 that would make the 55,000 number realistic. The actual number demobilized as of June 2007 (the time of fieldwork) was about 23,000; to date, the total is about 28,000, about half the 55,000 target proposed in 2004 This error is acknowledged in World Bank (2009).

World Bank's estimates for the sizes of the various forces is shown in Table 4.1. Both members of the national army and the rebel groups would be demobilized. DDR would occur in two phases, with the first phase (2004-5) involving 9,000 former rebels and 5,000 former army. Those not demobilized in the first phase would then join an interim integrated national army and police. Over the ensuing years, some 41,000 additional combatants from the newly integrated security forces were to be demobilized as part of a broader army and police reform program.

The DDR program in Burundi was part of the broader Multi-Country Demobilization and Reintegration Program (MDRP)[17], which was intended to embody a comprehensive strategy to "enhance prospects for stabilization and recovery in the region." The program documentation for the MDRP's programs reflected a strong presumption that demobilized combatants would likely suffer human and social capital deficits that would have to be remedied, otherwise "disaffected ex-combatants can pose a threat to stability." The MDRP characterized economic reintegration as the establishment of "sustainable livelihoods." Assistance was to do no more than would be "necessary to help them attain the general standard of living of the communities into which they reintegrate" (19). Social reintegration, in the form "reconciliation" between ex-combatants and civilians in their communities, was taken to be an important pre-condition for achieving a sustainable livelihood. The program took an "individual-oriented" approach—one that emphasized providing means to individual ex-combatants, as opposed to trying to do much at the level of communities; this orientation was apparently due to an implicit agreement that would have other agencies—namely UNDP—taking responsibility for community reconstruction.

[17]Participating countries included Angola, Burundi, Central African Republic, Democratic Republic of Congo, Rwanda, Republic of Congo, and Uganda.

4.4.2 The "treatment": reintegration program benefits

The reintegration program included a few components. The CNDRR directly administered a reinsertion subsistence allowance of between 30,000 to 185,000 Burundian Francs per month (FBu, with 500 FBu equivalent to 1 US 2007 dollar in PPP terms), depending on rank.[18] This assistance was provided over in 4 tranches over 18 months. Documentation from the program shows that at the time of our fieldwork in 2007, receipt of at least the first tranche was nearly universal among ex-combatants (98.5%).[19] Also, through offices set up in nearly all provinces and through "focal points" appointed in nearly all *communes*,[20] ex-combatants had access to various forms of counseling, including psychological counseling. These too were universally available. Thus, reinsertion allowances and counseling are not sources of variation in our study.

The source of variation that we study was the major benefit offered by the program: the so-called "socio-economic reintegration package." The package provided a menu of opportunities from which ex-combatants could choose. They could choose to be admitted into secondary school or university, to take up a one-year or shorter vocational-training program, or to take up a package to help begin "income-generating activities." The latter would involve working with one of the reintegration program's

[18]The benefits schedule was as follows: 31,446 FBU/month for infantry, 59,712 FBu/month for non-commissioned officers, 62,527 FBu/month for junior commissioned officers, 101,252 FBu/month for senior commissioned officers, and 184,754 FBu for top brass (CNDRR, 2005).

[19]Interestingly, the allowance was administered through transfers to special bank accounts for ex-combatants, giving many ex-combatants their first exposure to bureaucracy and the formal banking sector. Interviews in the field revealed that some ex-combatants applied what they learned through this by offering for-profit services to help other people deal with banks and government offices.

[20]Communes are the second-tier administrative units in Burundi, after provinces. A commune is a collection of villages and communities similar to a "county" in the western world. If we exclude the dense urban capital of Bujumbura, communes range in size from 15,000 to 100,000 individuals (or about 3,000 to 20,000 households).

implementing partners to devise a business plan, receiving basic training on running a business, and being provided with in-kind start-up materials for the new business (cash was not given). Regardless of the program chosen, the total value of benefits was to be 600,000 FBu or less. Program documentation shows that the "income-generating activities" option was by far the most popular among the ex-combatants that were online to receive benefits at the time of fieldwork in 2007: out of the just over 13,000 ex-combatants online to have received benefits by 2007, 96% took up the income-generating activities option, 3.6% took up vocational training, and less than 1% chose to resume formal education (MDRP, 2007). The income-generating activities assistance would often take the form of starting a small shop, a "moto-taxi" business, or buying agricultural assets to produce marketable farm goods. Thus, the 600,000 FBu in benefits would cover the cost of business counseling, any required permits, and the acquisition of materials, such materials to build a shop or kiosk, goods to sell, a motor-bike, a cow, or agricultural tools.

4.4.3 The source of exogenous variation: a disruption in access

The CNDRR and MDRP delegated the delivery of the socio-economic reintegration package benefits to NGO partners. Early in the program, when the number of beneficiaries processed and ready to receive benefits was rather small, a large and rather disorganized collection of local NGOs was contracted on an ad hoc basis to deliver benefits. In 2006, anticipating a surge in the number of ex-combatants who would be coming online to receive benefits, the MDRP decided to tighten up the system and contracted three large NGOs to deal with the coming wave of ex-combatants. These included Twitezimbere, a Burundian NGO, as well as PADCO and Africare, two international NGOs. The work was divided evenly among the three NGOs. PADCO was assigned to cover ex-combatants registered as residents in the south-west provinces, Twitezimbere was to cover ex-combatants registered as residents in the northern provinces, and Africare was to cover ex-combatants registered as residents in the center provinces. The assignments to the three NGOs were made by the end of the summer in 2006, and programming was to begin as soon as possible. The selection of Africare as one of the three partners was due to pressure by MDRP donors to use international NGOs as implementing partners; the pressure was due to certain procurement restrictions, it seems. This was the case despite the program administrators' and the CNDRR's concerns about the readiness of Africare to implement the program.[21] Indeed, it came to be a major problem: while PADCO and Twitezimbere were able to begin quite quickly, Africare's presence on the ground was barely established by late 2006.[22] This was followed by a contracting dispute between CNDRR and Africare that caused further delays. As a result, designated beneficiaries in the Africare area were denied access to the reintegration package until late 2007. This disruption in program access corresponded to the timing of our fieldwork: PADCO and Twitezimbere had begun reintegration programming by late 2006, whereas the problems with Africare's commencement of delivery would not be resolved until August 2007. Our fieldwork was conducted in June/July 2007. Thus, the respondents in our sample from the PADCO/Twitezimbere areas had access to reintegration programming, but those from the Africare areas did not, or at best, were only just beginning to have access. So long as any other differences between sample respondents from "center" areas and other areas can be controlled for, this bureaucratic breakdown provides us with a source of near-exogenous variation in program access. Such is the cornerstone of our strategy for identifying the effects of the reintegration program.

---

[21] Interview with Marcelo Fabre, MDRP, February 2009.

[22] Geenen (2007) notes that Africare had no presence on the ground in the area of her fieldwork, Ruyigi, in November-December 2006, and that they expressed concerns themselves about whether they could implement the program.

4.5 Causal identification

Two majors concern in the study of assistance programs such that is one are *self-selection* by beneficiaries into the program and *targeting* of beneficiaries by program managers. Self-selection and targeting are typically based on would-be beneficiaries' and program managers' predictions about how much beneficiaries will gain by participating. The conditions that inform such judgments are hidden to the analyst. When program access is universal, non-participants will tend to be those who would have little to gain from the program. Comparisons between those who had access and participated to those who had access but decided not to participate will tend to produce biased estimates of program effects. The bias is likely to be in the direction of the null hypothesis.

In our case, we avoid the problems of self-selection and targeting because we have a disruption in program *access* that had nothing to do with either beneficiaries' or program managers' choices. In the Burundi reintegration program, the disruption in program access was due to an idiosyncratic failure at a high level in the bureaucratic process. The arbitrariness of the disruption relative to the conditions of the beneficiaries considerably reduces the scope for bias due to self-selection or targeting.

Nonetheless, we cannot take this shock to program access as an unproblematic source of quasi-random assignment at the *individual* level. First, there may have been *incidental imbalances* in the characteristics of ex-combatants in the Africare versus non-Africare regions prior to the disruption to access. Also, there may have been differences in the types of communities in which would-be beneficiaries resided in the Africare and non-Africare regions. To reduce biases from these possible imbalances, we use covariate information to identify individuals in the non-Africare regions that resemble those in the Africare region in terms of their individual characteristics and the communities into which they are reintegrating.

Second, the disruption in program access occurred after some ex-combatants in the Africare region had already received a reintegration package during a preliminary phase of the program. An early cycle of the reintegration program was completed prior the decision to hand operations over to Africare and, therefore, prior to the disruption. Unfortunately, our data do not allow us to identify these individuals, although documentation from the program itself provides the true rate at which this occurred. The rest of the section explains our strategy for identifying causal effects in the face of these possible sources of bias.

4.5.1 Incidental imbalance

Because the shock to program access occurred at the region level (rather than the individual level), incidental imbalances in the attributes of ex-combatants and the types of home communities in the Africare and non-Africare regions threaten the validity of causal inferences drawn from comparing the two groups. In the analysis below, we study such possibilities by conducting balance checks on a rather large set of pre-disruption covariates. We find moderate to poor balance in covariates, as would be expected. Nonetheless, we have more than double the number of non-Africare recipients to Africare recipients. This creates the opportunity to construct a matched set of non-Africare respondents to use as a pseudo control group. To do this, we use a matching algorithm, GenMatch (Sekhon, n.d.; Diamond and Sekhon, 2005), to match Africare and non-Africare ex-combatants on fourteen covariates likely to be prognostic of economic and political outcomes. GenMatch uses an iterative "genetic search" which first specifies a modified Mahalonbis distance function, finds nearest neighbors, and then modifies the distance function to maximize covariate balance between treated and control units. For the balance metric, it uses the smallest p-value from t-tests (for binary variables) and Kolmogorov-Smirnov tests (for polytomous or continuous variables) of differences in distribution between treatment and control covariates. The process loops until it converges on a balance solution that cannot be improved-upon within tolerance limits. Many such loops are initiated, and the most balanced result is chosen. We match "with replacement," which means that we find the best match for each Africare subject even if that means using the same non-Africare subject more than once. This approach achieves the best covariate balance. Using these matched data, we estimate program effects by with a regression on an indicator for residence in the Africare region as well as all fourteen of the covariates used for matching. The inclusion of the covariates helps to reduce bias due to imperfect matching, while the matching helps to ensure that the model specification that we use for the covariates is of little import (Ho et al, 2007).

Given the rather modest sample size (less then 400 ex-rebel observations), matching achieves remarkably good balance on this large set of covariates. We cannot control for region-wide differences that distinguish the Africare region from the non-Africare region (i.e. region-wide "fixed effects"), as they would coincide perfectly with our shock to program access. An identifying assumption that we make is that once the incidental individual-level and community-level differences have been accounted for in our sample, there are no remaining region-wide differences. We think this is a reasonable assumption given the richness of our covariate set.

4.5.2 Exposure heterogeneity

The disruption in program access in the Africare region occurred after some reintegration programming had already begun. A subset of excombatants came online to receive reintegration assistance early in the program cycle (shortly after 2004). To prevent a large backlog from forming, the CNDRR commissioned a collection of small NGOs to begin administering reintegration packages at around the beginning of 2005. In the spring and summer of 2006, the CNDRR anticipated a large surge of ex-combatants coming online to receive assistance, and it was at this time that the program was centralized and contracts with the three major NGOs—Twitezimbere, PADCO, and Africare—were struck to handle the surge in cases.

Table 4.3 provides the relevant figures, obtained from the reintegration program itself. We see that as of the time of our fieldwork in July 2007, the program disruption affected 53% of designated ex-combatant beneficiaries in the Africare region. If we were to simply measure differences between Africare and non-Africare ex-combatants, we would obtain an estimate of the disruption effect that is biased toward zero. Such bias is attributable to the exposure heterogeneity among Africare-region ex-combatants. To correct for this source of bias, we use a strategy that weights our effect estimate by the inverse of the difference in disruption rates across the Africare and non-Africare regions. To see how this works, consider the following model,

$$y_i^* = \beta_0 + \beta_1 t_i + \mathbf{x}_i \beta_{\mathbf{2}} + \epsilon_i$$
$$t_i = \alpha + \phi z_i + \eta_i,$$

where y_i^* is a latent outcome variable, $t_i \in \{0, 1\}$ is an endogenous indicator of exposure to the disruption, $z_i \in \{0, 1\}$ and $\mathbf{x}_i' \in \mathbf{R}^k$ are exogenous variables in that $\mathrm{E}\left(\epsilon_i | (z_i, \mathbf{x}_i)\right) = 0$, and $\beta_{\mathbf{2}}$ is a $k \times 1$ vector. For the data considered here, the $z_i = 1$ refers to being in the Africare region, and $z_i = 0$ refers to being in the non-Africare region. The $t_i = 1$ refers to one's access to the program having been disrupted, $t_i = 0$ means access was not disrupted. The coefficient β_1 is our estimate of the program disruption effect. The coefficient ϕ measures the expected difference in the ex post probability of disruption for those with $z_i = 1$ relative to those with $z_i = 0$. The fact that t_i and z_i are binary variables does not affect our interpretation at all (see Angrist and Pischke, 2008, Ch. 4). We use a latent variable, y^*, for generality to generalized linear models that can be expressed in terms of latent variables (e.g. all binary, ordered, or multinomial probit/logit models; see Long, 1997). The transformations discussed in this section would then apply to differences on the latent variable rather than the observed response. When we are working with a linear model, then y^* simply equals the observed response.

We make the following exclusion restriction assumptions:

Assumption A.1: $\mathrm{E}\left(\eta_i | z_i\right) = 0$, and *Assumption A.2*: $\mathrm{E}\left(\epsilon_i | z_i\right) = 0$.

If A.1 and A.2 are true, then substituting the right-hand side of the expression for t_i into the right-hand side of the expression for y_i^*, we can write the the expression for y_i^* as follows:

$$y_i^* = \tilde{\beta}_0 + \beta_1(\phi z_i) + \mathbf{x}_i\beta_\mathbf{2} + \tilde{\epsilon}_i,$$

where $\tilde{\beta}_0 = \beta_0 + \beta_1\alpha_0$ and $\tilde{\epsilon}_i = \beta_1\eta_i + \epsilon_i$ with $\mathrm{E}\left(\tilde{\epsilon}_i|(z_i, \mathbf{x}_i)\right) = 0$. If we know the population value of ϕ exactly, then we can compute the ϕz_i values, and use the results of a regression of y^* on ϕz_i and $\mathbf{x_i}$ to obtain unbiased estimates of the effect, β_1, with the appropriate standard error.[23]

Documentation from the program (MDRP) itself allow us to obtain a value for the true rates of disruption in the Africare and non-Africare regions. We assume that these rates are equivalent to the ex post expected probability of disruption conditional on whether one was located in the Africare region or not. Our discussion with a key program official suggests that this is a reasonable interpretation, as pre-disruption access seemed to be quite random.[24] If one is wary about this, one may take a conservative position, and assume any estimates based on this interpretation form an upper bound; a lower bound is available by not correcting for ϕ at all, so informative bounds can be constructed.

Table 4.3 shows the rates of disruption for the population of demobilized combatants registered to receive benefits in the Africare and non-Africare regions. Based on this information, we take $.53 - 0 = .53$ as the value of ϕ. The MDRP documentation also provides information on the rate at which beneficiaries who were due to received benefits after the NGO transition actually *took up* the program by the time of fieldwork in July 2007. These rates reflect the disruption in the Africare region: at the time of fieldwork, about 66% of these post-transition beneficiaries in

[23]See Murphy and Topel (1985) for properties of estimators such as this that rely on obtaining "missing" regressors from auxiliary information. In our case, since we are assuming that we know ϕ exactly, there is no added approximation error, and thus no additional inflation of the standard errors. Another way to look at this is as a two-sample instrumental variable estimator, along the lines of Angrist and Pischke (2009:147-150), but where the first stage parameters are known exactly.

[24]Interview with Marcelo Fabre, MDRP executive offices, February 23, 2009.

non-Africare regions had already begun their engagement with the program, while only 16% had begun to do so in the Africare region. We stress that we *do not* want to incorporate this information into the construction of ϕ. The rate of program take-up in the non-Africare region reflects, to a large extent, self-selection; those who had not taken up the program by then should nonetheless be considered as having had uninterrupted access. Individuals in the non-Africare region should have factored the availability of the reintegration program into their decision-making, and so the effect of access to the program should be evident even among those who have chosen not to take it up. The very small number of beneficiaries having begun engagement with the program in Africare regions after the disruption had only recently done so at the time of fieldwork. Considering them to have been subject to the disruption has, at worst, an effect of biasing our effect estimate toward zero. Thus, to be clear, we are measuring the effects of *disrupted access* to a functioning reintegration program. By disregarding these rates of program take-up, we actually take a more conservative approach, in that we do not attempt to make the program look more effective by netting out any dilution due to non-engagement by certain beneficiaries despite their having had access.

The above results are based on a homogenous effects model. If we loosen that assumption to allow for variable coefficients, then we could interpret our estimate of β_1 as the effect of the program only on those whose access to the program would depend on z_i (see Angrist and Pischke, 2009, Theorem 4.4.1). It seems to us that the estimate under the homogenous effects assumption is probably higher than the true average effect of the program. This is because people who obtained access to the program in the Africare region are likely to have been more organized generally about their personal affairs; therefore, we suspect for them, the program's benefits would be less. For these reasons, we believe that the scaled estimates are likely to be an upper bound on the true average effect. Reduced form estimates—that, is estimates from regressions using the unadjusted z_i rather than ϕz_i—are a lower bound. We report both below.

4.6 Data

The ex-combatant data are drawn from the multi-purpose *Wartime and Post-Conflict Experiences in Burundi* survey.[25] The survey includes data from interviews with civilians, as well as ex-rebels and ex-army, both demobilized and those integrated into the new security forces. This paper works with only the demobilized ex-rebel data. The data were obtained through a multistage random sample from lists of demobilized ex-rebels registered to receive reintegration benefits through Burundi's national DDR program (the CNDRR). The first stage involved randomly selecting half of Burundi's 129 *communes*—Burundi's second-tier administrative unit.[26] We then set a target number of ex-combatant interviews to complete in each commune, with targets proportional to the national proportion of ex-combatants registered with the DDR program in the commune. Targets ranged from 2 to 33. Then, we obtained from the national DDR office the complete lists of ex-combatants registered as residents of each of the selected communes. We then drew a simple random sample (with a random number generator in the software package, R) of the desired number of interviews from each of these commune lists. We created a randomly selected reserve list for each commune to use in the case of non-response. Selected participants were contacted and brought to the respective Provincial Bureau by DDR program staff for interview on a scheduled date by our enumeration team. The rate at which our first choice was interviewed was very high—about 90%—and so we assume no need for further adjustment to account for non-response.[27] This very high response rate was

[25]Details on the survey are available at www.columbia.edu/∼cds81/burundisurvey/

[26]See fn. 20.

[27]Keep in mind that the covariate adjustment that we perform below will reduce any biases associated with those factors. With such a low non-response rate, it is reasonable to believe that any residual bias is negligible.

likely due to a few factors: (1) respondents probably took the interview to be a requirement of the DDR program given that they were contacted by CNDRR program staff themselves; (2) we accommodated respondents' schedules by setting dates for interviews well in advance; (3) the fieldwork was conducted during the idle interim period between planting and harvesting seasons, and so respondents working in the agricultural sector faced few competing demands on their time; and (4) while participation was voluntary, a "transport allowance" of about 2 US dollars was provided to each respondent after they completed the interview, thus making it worthwhile for respondents to sit through the entire interview. With the observations limited to former rebel males registered to receive reintegration assistance outside Bujumbura, the dataset includes 110 ex-rebels registered in the Africare region and 261 outside the Africare region.

As is common in large scale surveys such as this, the data exhibit occasional missingness on items. Such missingness is due to "don't know" responses, enumerator error, or data-entry error. In the subset of data that we used, only once did the item missingness rate reach 10% for any given variable—such was the case for the measure of the death rate in an ex-combatant's fighting unit; otherwise the rates of item missingness were usually 0, but occasionally around 1-2%. Listwise deletion would nonetheless have implied discarding 23% of excombatant observations. To avoid the inefficiencies and biases associated with listwise deletion, we used multiple imputation to fill missing values.[28] Multiple imputation was conducted using the "mice" package

[28]In additional to the loss in statistical power, listwise deletion can bias results by either discarding cases based on the value of the dependent variable or distorting the relationship between the sample and population in ways that bias approximations (e.g. linear or quadratic) of unknown functional forms. This is explained in Samii (2010). If multiple imputation is done in a reasonable manner, it will be considerably less biased and will generate appropriate uncertainty estimates. See King et al (2001) for some general discussion of multiple imputation methods, and Rubin (1987) and Little and Rubin (2002) for deeper treatments. Approximation bias is best understood through the lens of non-random sampling. See, e.g., Korn and Graubard (1999:159-185).

for the R statistical computing environment (van Buuren and Groothuis-Oudshoorn, 2010). "Predictive mean matching" imputation was used for all numeric, binary, and ordered categorical variables; multinomial logit regression imputation was used for non-ordered categorical variables. Predictive mean matching (also known as "nearest neighbor hot deck") is attractive because, when the item-level data exhibit only low levels of missingness, it is robust to a wide range of misspecifications of the imputation model (Little and Rubin, 2002:69). Uncertainty in predictive mean matching arises in the fit of the predictive mean model. With the multinomial logit, uncertainty arises from the fit of the model and the fundamental uncertainty inherent in the outcome being modeled as a random draw from a multinomial distribution. To account for these imputation uncertainties, multiple imputed datasets should be generated (hence "multiple" imputation). The analysis should then be run on each of the datasets, and averaged using the appropriate averaging forumlas. This was the approach that we took, working with five imputed datasets.[29] Summary statistics for the five imputation-completed samples are presented in Table 4.4. The geographic distribution of respondents is displayed in Figure 4.1. We discuss this summary information in detail in the section below on covariate balance.

4.7 Inference

For inference, we take covariate and treatment values as fixed and base our inference on presumed sampling variability. On each of the imputation-completed datasets, we use sandwich estimates of coefficient covariances that account for clustering at the *commune*-level. We consider communes to be the relevant area within which an individual's routine economic interactions is confined.[30] For the quantile regression

[29]As Rubin (1987:114) explains, rather few imputations are typically needed to achieve reliable estimates.

[30]Refer to fn. 20 for an explanation of *communes*.

estimates, we use standard errors from inverted rank-test confidence intervals asymptotically robust to non-iid sampling (Koenker and Hallock, 2000).[31] To combine the estimates from each of the imputation-completed datasets, we used standard formulas based on the properties of mixed normal distributions (known as "Rubin's rules"; Rubin, 1987; King et al, 2001:53). Estimation was conducted in both R and Stata. In Stata, when possible, we used the imputation combination function provided by the Clarify software (Tomz et al, 2003). When this was not an option in Stata, and for the estimation in R, we used our own imputation combination programs.[32]

4.8 Covariate Balance and Matching

4.8.1 Covariates

A covariate should be considered confounding if it is (1) correlated with the treatment, (2) correlated with the outcome of interest, and (3) temporally prior to the treatment. We made the case above that variation in our treatment (in this case having program benefits withheld in the areas serviced by Africare) is due largely to an exogenous shock. In other words requirement number (1) above is not due to self-selection or targeting of program beneficiaries. However as discussed above, our treatment is not truly the result of a randomized experiment and incidental imbalances may exist. Ex-combatants in the Africare areas might be different according to important covariates than their fellows in the other areas of the country. Therefore we have to be concerned about making inferences from extreme counterfactuals (King and Zeng, 2006). Since we do not know the true functional form of the relationship between our covariates and our outcomes of interest we cannot assume that we know the relationship between a given independent variable over the whole range of a

[31] Bootstrapped confidence intervals are inappropriate on the matched sample (Abadie and Imbens, 2008).

[32] The programs are available at http://www.columbia.edu/∼cds81/.

given the independent variable. By fitting a particular function (say linear) for to the relationship between independent and dependent variables we are implicitly assuming that the relationship between a given independent variable was the same for very high levels of that variable and very low levels of that variable. Pre-processing the data with matching overcomes this problem by making comparisons between treated and control cases that are very close to each other in terms of the values of their covariates. This substantially reduces the implications of our functional form assumptions (Ho et. al, 2007; King and Zeng, 2006)

In what follows we list the variables on which we matched and our reasons for using them. We divide the set between covariates that measure (i) individual characteristics versus (ii) community characteristics. We need both types of covariates, because economic outcomes are a product of one's individual characteristics interacting with one's economic environment. Community characteristics were measured at the level of commune.[33] There is no concern with post-treatment bias by including these commune-level covariates, because ex-combatants chose their locale of residence and registered it with the reintegration program before the disruption in the Africare region. Accordingly, our sample frame listed ex-combatants by the commune where they registered and this is what we assign to the subjects in our study. As described below, our study is limited to ex-combatants residing outside the capital, Bujumbura. Communes outside the capital range in size from 3,000 to 20,000 households, meaning that they are quite large and they generally circumscribe the domain of an individual's routine economic affairs.

4.8.1.1 Individual characteristics

Age We hypothesize that younger ex-combatants will be less well endowed with pre-existing skills and social networks that are necessary for successful economic reintegration. Previous studies (Blattman and Annan, 2009) have shown that

[33] See fn. 20 for more information on communes.

child soldiers in particular are deprived of the crucial period of their lives when important human and social capital skills are developed. Thus or expectation is that younger ex-combatants may experience less economic reintegration than older ex-combatants.

Hutu We expect differences in the earning potential for those identifying as Hutu versus Tutsi (there are no Twa in this sample). The vast majority of those participating in the rebellion were Hutu. The small number of Tutsi that participated in the rebellion are likely those whose relations to their communities were somehow strained. For those reasons, we expect Tutsi ex-rebels to fair worse economically than their Hutu counterparts.

Father in Agriculture Prewar An ex-combatant coming from a pre-war farming household was probably relatively poor compared to those ex-combatants whose parents had more lucrative non-agricultural jobs. Coming from a non-farming household is also likely to affect an ex-combatant's ability to find an occupation outside of farming.

Years of Education, Pre-War Ex-combatants with more education should have an easier time finding a productive livelihood and reintegrating.

Prewar Household Wealth Index Because ex-combatants with greater wealth before the war should, we presume, have higher incomes after the war we include a scale that captures the value of an ex-combatant's pre-war household assets. We created the scale based on survey respondents' responses to three questions about their pre-war household: did their beds have sheets, did they own a radio, and did they own any cattle. The bed-sheets and radio measures were devised after focus groups in Burundi, in which these assets were deemed clear markers of a household having moved beyond basic subsistence in their consumption level. Ownership of cows is a distinct indicator of wealth in Burundian society.

We fit a two-parameter logistic graded response model on these responses, using the "ltm" package in R (Rizopoulous, 2006). We then generated a pre-war wealth index using factor scores from this model, estimated with the imputation method described in Rizopoulos and Moustaki (2008).[34] We fit the model on a pooled dataset that included the ex-rebels and civilians from the same regions as the ex-rebels. We assume that the factor loadings are common across the two groups; including the civilians greatly improves the precision of the fit.

Out-going Rank Outgoing rank is likely to indicate the ability of the individual. In addition, outgoing rank determined the monthly reinsertion benefit level.[35] We use two binary variables to indicate whether the subject's exiting rank was commissioned officer or non-commissioned officer, with rank-and-file as a baseline category.

Unit Death Rate We are concerned that ex-combatants who experienced the most mortal combat would be more psychologically traumatized and therefore might find it more difficult to return to normal economic activity. Thus we include the rate of death within subject's units as reported by the subjects.

Years in Faction We hypothesize that ex-combatants who were in combat longer would find it more difficult to reintegrate into society. In Uganda, Blattman and Annan (2009) found that time in the LRA faction was associated with poorer economic reintegration. Their explanation was that time in the faction robbed

[34] We also asked respondents about whether they owned land before the war, but we did not include it in the scale because there was too little variation in the variable (about 95% of respondents indicated that they possessed land prior to the war). A Mokken analysis of these three indicators showed that they loaded monotonically and contributed a moderately strong, single dimension signal (the lowest H score was 0.38; see van der Ark, 2007). We used ANOVA tests to determine whether a one, two, or even higher order logistic model fit the data best, and found that the two parameter model was optimal.

[35] Refer to fn. 18.

individuals from having time to develop economically useful capabilities and networks.

Family Death Rate Our concern was, again, that more exposure to violence would make economic reintegration more difficult. With this covariate we capture exposure to violence in terms of proportion of family members killed in the war. Furthermore postwar economic opportunities are often supplied by social networks like the family. Ex-combatants with a large percentage of family members killed would, we speculate, have fewer contacts with which to pursue employment.

Non-CNDD-FDD Faction Member While several rebel factions fought in the war the CNDD-FDD came out the clear winner of the conflict. As such CNDD-FDD members may have better access to economic opportunities than members of other factions, like the CNDD-Nyangoma, FROLINA, or FNL. Therefore, we expect CNDD-FDD members to have a more successful economic reintegration experience than non-CNDD-FDD members. In addition, the factions differed in there levels of internal discipline. FROLINA and FNL were associated with more abusive tactics, and so control for this faction helps to address Humphreys and Weinstein's (2007) finding that more abusive factions have more problems with social acceptance.

Demobilization Date Ex-combatants who demobilized earlier had a longer period to reintegrate than did those who demobilized later so we expect ex-combatants who demobilized later to have less success in economic reintegration. We use the number of months that elapsed from the start of demobilization to the time of the subject's demobilization, standardizing the variable to have mean 0 and standard deviation 1 to improve numerical stability.

Propensity Score In order to insure overall balance of our covariates we also matched on the propensity score. We estimated the propensity score using a logistic regression of Africare region status on all of the covariates (including the community characteristics mentioned below). The specification did not include any interactions or higher order terms.

4.8.1.2 Community characteristics

Accumulated War Violence, 1993-2004 We hypothesize that communities that were exposed to more violence may have suffered capital destruction that dampens economic opportunities. The accumulated war violence for each commune is the sum of Gaussian-kernel distances between the commune's geographic centroid and each of the major conflict events listed in the Armed Conflict Location and Event Database (ACLED, Version 1), produced by the Peace Research Institute of Oslo.[36] We use the sum of kernel distances, rather than simple counts, to account for the fact that the ACLED data likely under-reports the number of violent war events, and that such events tend to occur in geographic clusters that spill over commune boundaries.

Log(2004 Population density) Population density is closely associated with the economic vibrancy of a commune. More densely populated communes will have larger markets for products and labor. We use the natural log to account for extreme skew in the distribution.[37]

Ruling-Party-controlled Province Provinces with governors affiliated with the ruling CNDD-FDD party would likely be more successful in attracting government-controlled funds for reconstruction, which is an economic stimulant. We assign

[36] The dataset is available at http://www.prio.no/CSCW/Datasets/

[37] Extreme skewness and other departures from ellisoidal distributions make it difficult for any matching algorithm to ensure that improvement in the balance over any single covariate does not introduce serious imbalance over other covariates (Rubin, 1976; Diamond and Sekhon, 2005). We also think the log-transformed value is better for characterizing the distance between subjects.

a dummy variable for whether the subject's commune is located in a province with a CNDD-FDD governor.

Bujumbura Bujumbura is the capital city and, given its enormous size and density relative to the rest of the country, it is clearly a case apart in terms of potential economic outcomes. Because the Africare region is outside Bujumbura, we simply exclude ex-combatants from Bujumbura entirely from the analysis. Thus, the inferences we make below refer are based only on ex-combatants residing outside Bujumbura.

4.8.2 Balance

Balance statistics for the five imputation-completed datasets are presented in Table 4.4. As discussed above, we used matching with replacement, and multiple Africare respondents were matched to common non-Africare respondents in a number of cases. This is evident from the sample sizes reported in the first column of the table. We see that out of the 261 candidate subjects in the non-Africare region, between 63 and 75 were selected as "controls" across the five datasets. In that way, the matching algorithm was quite selective, relegating nearly 200 of the non-Africare region subjects as being too distant from the Africare subjects to be useful for comparison. As recommended in Ho et al (2009:16), we ensure that each of the matched non-Africare observations has the appropriate influence on our analysis by weighting each of them proportionally to the number of times each is used as a match.[38] The balance checks incorporate these weights. We also use these weights in our estimation of program effects below.

---

[38]The precise weight is $k_i \frac{N_c}{\sum_i k_i}$, where k_i is the number of times that i is matched and N_c is the number of control observations. The normalizing constant, $\frac{N_c}{\sum_i k_i}$, simply ensures that the weights sum to the number of controls in the sample and does not affect the analysis. These are the weights produced by default in the R package, MatchIt (Ho et al, 2009), which we used to implement GenMatch.

Some of the covariates are well-balanced even without matching. This was the case for age, Hutu, prewar education, non-commissioned officer, demobilization date, and ruling party province. We nonetheless match on them to insure that they remain in balance even after matching on the other covariates that were not naturally in balance. The fact that several variables were balanced before matching has implication for how we interpret the balance statistics after matching. Ho et al recommend looking at improvement in balance rather than p-scores on t- and KS tests. Their rationale is that many observations are dropped as a result of matching procedures and so p-scores can improve simply by virtue of the standard errors increasing as the sample size gets smaller. Imai et al (2008) recommend, wisely we think, that the researcher also make sure that balance is actually improving—e.g., means are getting closer together and observations are getting closer to the 45 degree line in the QQ plots. We agree with their argument but hasten to point out that it is not necessarily cause for concern that some covariates' balance worsens as we attempt to achieve balance on other covariates so long the resulting dataset does not exhibit serious imbalance. For this reason, we do not find the occasional losses in balance on certain covariates (e.g. on age, Hutu, or education) to be reason for concern. In addition, even though t-test and Kolmogorov-Smirnov (KS) p-values do not appropriately reflect certainty about equality of distributions, they remain useful as standardized measures of imbalance (Diamond and Sekhon, 2005:11).

The statistics in Table 4.4 show that we achieved very good balance after matching, even with this large set of covariates. The t-tests on difference of means indicate that the distribution of our treated and control cases have very similar means, with only log(population density) exhibiting somewhat low p-values (e.g. below .20). We inspected density plots for this variable in the matched data, and found that the Africare respondent communities exhibited bimodality on this variable, while the non-Africare respondent communities were more uniformly distributed. Thus, the imbalance was mostly on interior values. This is not of major concern, because regression adjustment tends not to be very model-dependent when imbalance occurs over interior regions (that is, within the convex hull) of the data (King and Zeng, 2006).

The Kolmogorov-Smirnov (KS) tests provide a measure of balance on higher order features of covariate distributions. The tests show that we achieve excellent balance for about half of the ten variables where the KS test was appropriate. Among those with rather low p-values, we find that the t-test p-values tend to be large, and the ratios of treated and control variances tend to be reasonably close to one over the imputation-completed datasets. Thus, we have balance on at least the first two moments. One variable, the wartime violence index, tends to exhibit rather high treated/control variance ratios. We inspected density plots for this variable in the matched data. We found that the problems are attributable to a small mass of 8 Africare respondents with values located just above the maximum of the non-Africare respondent values. This means that we have to accept some degree of extrapolation in this range.

Our estimation strategy is to specify a regression model that uses a linear specification for all covariates. Regression on the matched data provides counter-factual values at precisely the covariate locations of the Africare observations, helping to overcome the fact that we cannot matched exactly on all covariates (Rubin and Thomas, 2000). The balance that we have achieved on the covariates substantially reduces the import of our choice of a linear model. The covariance adjustment also makes the analysis more efficient to the extent that the covariates are prognostic. Such efficiency is important in our case because we have a small sample size.

4.9 Programmatic impacts: economic reintegration

We now present estimates of program effects. We begin with effects on economic reintegration. As discussed above, these are estimates of *programmatic effects* in that the reintegration program provided direct assistance to ex-combatants for realizing a productive civilian livelihood. From an academic perspective, effects on economic reintegration are a necessary part of the causal chain outlined above, which proposes that a boost to one's economic conditions can induce an individual to subscribe more faithfully to the laws and norms that govern civilian society. Our estimates of economic effects provide a basis on which we can judge whether these data provide a good test of this proposition. If we do find strong economic effects, then we are in a good position to assess the claim that political integration follows from a boost to economic well-being. Of course, for well-understood reasons associated with the study of mediated effects, this kind of analysis can only test the plausibility of such a two-step causal process (Green et al, 2010).

From a policy perspective, it might seem trivial that there should be *some* effect on livelihoods, but remarkably, past studies in Sierra Leone and Liberia (e.g. Humphreys and Weinstein, 2005; Levely, 2010) raise questions about whether such benefits of reintegration programs are manifest. If we also find that they are not, then the implications are that some serious reconsideration needs to go into the design of reintegration assistance. For example, it might strengthen the case made by some that individual-level targeting is ineffective without strong community-level assistance (Muggah, 2009). If we do find substantial effects, then it raises question as to whether conditions in Burundi were different than in Sierra Leone or Liberia, or whether the latter studies generated biased estimates. As we discussed above, there are good reasons to suspect the latter.

We present results for two measures of economic reintegration: (1) monthly income and, as a consequence of income, poverty incidence; and (2) livelihood, specifically the nature of the occupation that the respondent obtained. We present unadjusted estimates on the raw data and then regression adjusted estimates on the matched data. Having the two side by side helps us to assess the bias reduction as well as potential efficiency gains and losses from the adjustment methods. Our primary estimates of average individual-level effects are the regression-adjusted estimates on the matched data. As discussed above, a lower bound treatment effect estimate comes from regressions that do not correct for exposure heterogeneity, and an upper bound coming from regressions that do. We also present graphs that display estimated causal effects for the population of ex-combatants serviced by Africare. These show what would have been the gross impact of the program on the Africare beneficiaries had there been no interruption (that is, the "effect of the treatment on the controls").

4.9.1 Income and poverty incidence

Our outcome measure of income is the natural logarithm of monthly personal income (in Burundian Francs, or FBu) reported by the responded, with 1 FBu added. We used the natural logarithm as a variance stabilizing transformation to reduce estimation sensitivity to outliers. We added 1 FBu to handle the few cases of zero reported income. To check for sensitivity associated with adding 1, we also fit a Tobit regression on log of income with zero-income observations treated as censored; results were identical and so we do not display them. Inspection of the distribution of log monthly income variable over the covariates showed that the transformation was effective in removing heteroskedasticity, although the overall variance of the logged variable was still quite high.

The left panel in Figure 4.2 shows the income distributions for Africare respondents and the matched controls from the five imputation-completed datasets. The gray blocks in the plots show regions of the income distribution below the $1.25/day poverty line (at purchasing power parity, equivalent to 15,000 FBu/month). The dots for the matched controls are scaled proportional to the weight they receive in the analysis. We see that the Africare distribution is substantially heavier at lower income values (e.g., below 10,000 FBu/month). Indeed, the most striking visual impression is that the matched control income distribution exhibits a "floor" just below 10,000 FBU/month, whereas this floor is not evident for the Africare respondents. The Africare respondents also include a cluster of high earners (centered on 100,000 FBU/month) in an area where the matched control distribution is very thin.

We quantify these differences by examining differences in means, with linear regression, and then differences in quantiles, with quantile regression. The linear regression estimates appear in Table 4.5. The results show that in the raw data, Africare respondents' monthly income was about 50% lower on average. On the matched data, before adjusting for exposure heterogeneity, the point estimate for the Africare coefficient very similar as in the raw data (-0.59 versus -0.67), although the p-value (0.26) is quite large for the matched data. The stability of the coefficient estimate suggests there was little confounding to be removed in this case; the inflation of the p-value reflects the rather small sample size in the matched data. When we run a leaner specification that includes only the unit death rate, war violence index, ruling party province, and the propensity score, the estimate does not change much (coefficient of -0.53 with p-value of 0.29). The next estimate in the table adjusts further for exposure heterogeneity, providing our upper bound estimate. The adjustment essentially scales the Africare coefficient by about 1.9. This estimate is suggests that Africare respondent incomes are 70% lower income on average, but with the same large p-value. Thus, we find only very weak evidence that reintegration programming boosts incomes *on average* in the short run.

Differences-in-averages often hide important differences between groups' incomes, because income distributions can be very poorly behaved in terms of skew or massing at extreme values even after stabilizing transformations (Koenker and Hallock 2000; Angrist and Pischke, 2008:269-270). We have already seen visually that there appear to be important differences in the income distributions of the Africare respondents and their matched counter-parts. We used quantile regression to estimate more general income effects, including effects on incidence of very low or very high incomes. We estimated differences in income deciles 1 through 9. The estimates are shown in Table 4.6. We see large differences in the decile values, and p-values are quite small (given the sample size). Thus, while there is essentially equality of distributions for the first two deciles, the from the matched data we estimate that the third decile among program non-recipients in the Africare region is 45% (lower bound) to 68% (upper bound) than what it would have been had there been no interruption. Interestingly, these differences are decreasing over the deciles. By decile 7, the gap narrows to a difference of 32% (lower bound) to 52% (upper bound). Thus, we find that the income effect is declining in potential income, at least on the log-scale. One reason for this pattern could be a statistical artifact: if the income effect was, for the most part, a simple linear shift, then it would appear as a declining percentage shift after the log transformation. Running the same linear regressions on the raw (not log-transformed) income measure, we estimate that Africare respondent incomes were between 7050 (lower bound) to 13,300 (upper bound) FBu/month lower on average, although the p-value again is high at 0.17. Another explanation may be that the pattern is a genuine reflection of diminishing returns of the program over potential incomes. The data do not allow us to distinguish between these two stories; it is reasonable to believe that some combination of the two is at work.

The quantile effects are easier to understand when they are used to plot cumulative log-income distributions. We do so in the right panel of Figure 4.2. The results indicate dramatic effects on poverty incidence. The actual cumulative log-income distribution for Africare-region respondents is plotted with the black dots, and fitted values from the quantile regressions are given by the black line. Then, we plots the lower bound (white dots) and upper bound (gray dots) estimates of what the Africare respondents' cumulative income distribution *would have been* had there been no program disruption. The gray area shows the range of incomes below the poverty line. Poverty incidence among the Africare respondents is shown to be about 60% (the point where the black line crosses from the gray into the white region). The counterfactual distributions show that this would be an estimated 20 percentage points (lower bound estimate) to 35 percentage points lower had there been no disruption. We can see that this difference is not sensitive to the precise location of the poverty line. A regression on a poverty incidence indicator (not shown) estimates precisely the same effect with a p-value of 0.02.

What is the significance of this income effect relative to the value of the program benefits? An answer helps to assess the material benefits of the program against the costs. Above, we described that the the value of the benefit was up to 600,000 FBu. That value was allocated in the form of training, fixed or durable assets (e.g. materials to build a shop, some cows, a scooter, etc.) and immediately sellable assets (e.g. products, like beer or Fanta, to sell at a new shop). From the cumulative log-income distribution plot in Figure 4.2, we see that the median earner in the Africare region was earning about 10,000FBu/month. We estimate that this is about 64% (lower bound) to 44% (upper bound) of what that person would have earned had there been no disruption, implying a counter-factual incomes of about 15,500 and 23,000 FBu/month. Thus, the full value of the program benefit could not be immediately converted into income. Rather, the income effects 5,500 to 13,000 FBu/month imply that it would take 4 to 9 years to recuperate the value of the 600,000 FBu investment. This recuperation depends on the short-run income differences that we measure being sustained over such a period—something that we cannot assess. Future research might the factors that constrain ex-combatants' ability to convert program benefits into income.

4.9.2 Livelihood

To study effects on livelihood, we examine respondents answers to a question about what kind of work they did to sustain themselves. We coded their responses as (1) no occupation, (2) an agricultural occupation, (3) a non-agricultural, skilled sector occupation (including returning to school, professional position, or skilled labor, such as automotive repair, electrical technician, etc.) and (4) or a non-agricultural, unskilled sector occupation (including security guard, manual labor, etc.). The distribution for the Africare respondents and their matched controls are displayed in Figure 4.3 (the displays average over the five imputation-completed datasets). We see that agricultural occupations dominate in a major way, accounting for about 60-70% of responses overall. We see some indication that Africare respondents are more likely to report that they have "no occupation" and less likely to report a "skilled" occupation. Looking more closely at these differences, we found that for "skilled" occupations, the pattern is driven by differences in attainment of skilled labor occupations, rather than returning to school or obtaining a professional position. Indeed, there were *no* Africare respondents reporting having obtained a skilled labor occupation, whereas the rate among the matched non-Africare respondents ranged from 7% to 9%.

To quantify these differences, we fit a multinomial logistic regression of the occupation categories, with agriculture set as the baseline category. The regression estimates appear in the upper panel in Table 4.7. In the raw data, an Africare respondents' relative likelihood of a job in the skilled sector versus in agriculture was 63% lower than for a non-Africare peer; the association is strong (p-value = 0.03). Associations for the other categories are not strong as indicated by the high p-values. In the matched data, the point estimate for this effect remains about the same (relative risk ratio of 0.39 versus 0.37), but the p-value is quite large (0.23). We fit a model with a leaner covariate set, only using covariates with p-values less then 0.10 in the full specification—this included commissioned officer, unit death rate, war violence index, and log(pop. density). The results barely changed. Thus, we have only very weak evidence that the program changed the distribution of ex-combatants over livelihood types. We plotted the gross effects on the distribution of livelihoods in the right panel of Figure 4.3. The black dots are based on the fitted values, and the white and gray dots show estimated lower bound and upper bound counter-factual values, respectively. The estimates suggest that the rate of skilled occupation attainment may have been around 5 to 10 percentage points higher had there been no interruption, with compensating decreases coming mostly from the agricultural sector and the segment with no occupation. But given the very low levels of significance, these estimates should be considered very rough. It is clear that the strong tendency to seek agricultural livelihoods overwhelms any potential program effect on occupations.

While there is little to suggest a fundamental transformation in livelihood outcomes, to what extent can we say that livelihood outcomes generally *improved*? In order to measure this, we need to impose some ordering on the livelihood outcomes. Based on our observations in the field and discussions with Burundians, we felt the following ordering made the most sense: "no occupation" would be the least desirable category in terms of social status and ability to sustain oneself, "agriculture" and "non-agriculture, unskilled" occupations would be at the same level of desirability, and "skilled" occupations would be to be the highest. We fit an ordered logistic regression to these outcomes on the raw data and then the matched data, using the same covariate and exposure heterogeneity adjustments as in the above regressions. From the raw data, we see that overall Africare respondents have 20% lower odds of being in a higher ranked category than non-Africare respondents. When we remove confounding bias through matching and regression adjustment, the difference becomes much larger. Indeed, we estimate that the effect of the reintegration program was such that Africare respondents had between 54% to 77% lower odds of being in a higher ranked income category. The p-value for this difference is 0.13, which given the sample size provides moderate evidence of a genuine difference between the two groups.

4.10 Downstream impacts: political reintegration

To what extent do programmatic economic impacts contribute to political reintegration? That is, can a boost to one's economic well-being affect one's orientation toward political order, deepening one's appreciation of norms that govern civilian society? The empirical analysis thus far shows that the reintegration program substantially boosted the income of ex-combatants whose earnings would otherwise have been low. Effects on those with higher potential income were not statistically significant by any discriminating standard. Nonetheless, the conditional income effects resulted in large differences in poverty. These kinds of concentrated effects may be desirable. The logic of opportunity costs suggests that it is precisely these low potential-income ex-combatants who may be most at risk to be induced into action against the state or crime. Concentrating benefits among this segment may be the most efficient way to contribute to post-conflict stability. We find moderate evidence of effects on livelihood outcomes, with program recipients being somewhat more likely to have a more desirable occupation. However, these moderate effects do not result in any substantial change in the overall distribution of ex-combatants over livelihoods.

As described above, we define political reintegration as an ex-combatant accepting that violence is not a legitimate means to seek political or personal gain. Such reintegration may be reflected in attitudinal changes. These include a favorable view of civilian life relative to combatant life and attitudes consistent with norms of peaceful and democratic conflict management. Such reintegration should also be reflected in conduct—e.g., in refraining from participation in political violence or using threats for personal gain. Ideally, one would have measures of both attitudes and conduct. Unfortunately, our study did not acquire such behavioral measures, and so we can only examine effects on the expressed political attitudes. Attitudinal questions such as these tend to be quite noisy. Thus, we expect this to be a rather low powered test of political reintegration effects. With those caveats, we consider three attitudinal measures: (1) responses to a question about whether life was better as a combatant; (2) responses to a question of whether one is very satisfied, somewhat satisfied, or unsatisfied with the peace accords, which were ostensibly based on norms of peaceful and democratic political competition; and (3) responses to a question about whether Burundians should allow the government as currently constituted to have a lot of time to sort out the many problems facing Burundian society, or whether Burundians should start thinking about changing the government.

In studying down-stream political reintegration impacts, we actually focus on reduced form relationships between Africare status and these outcomes. This may seem like a strange way to test what is essentially a mediated relationship between a reintegration program and political reintegration, with economic well-being being the mediator. The reasons have to do with what our quasi-experiment allow us to identify. We cannot use Africare status as an instrumental variable for any single economic outcome, because we have already seen that Africare status is associated with substantial differences on multiple dimensions—even if the statistical significance of some of these differences is very low. Recently expounded problems with linear mediation analysis are so profound (Green et al, 2010) that we see little added value from such an analysis. Nonetheless, we think our reduced form estimates provide a useful probe into the plausibility of the claim that economic reintegration contributes to political reintegration. The reintegration program in Burundi was tightly focused on individual-level economic assistance. Thus, we see little reason to doubt that any apparent effects on political reintegration arise through the economic assistance role of the program.

4.10.1 Preference for civilian life over combatant life

We asked whether respondents thought life is better as a civilian (=1), with the alternative (=0) being that life was either better as a combatant or there was no difference. About 81% of matched controls indicated that life was better as a civilian life, whereas only 71% of Africare respondents indicated as such, implying 43% lower odds for Africare respondents. Thus, the strong tendency overall is to indicate a preference for civilians life. In depth interviews that we conducted with ex-combatants alongside the survey, we were told about a general sense of "war fatigue." This is evident in the data.

More refined estimates of the program effects are displayed in Table 4.8, which present results from logistic regressions. In the raw data we see that the Africare coefficient is essentially zero. In the matched data however, we estimate that the odds of stating that civilian life is better are 51% to 74% less among Africare recipients. However, the p-value of 0.17 is large, indicative of imprecisely we have measured the effect. None of the other covariates included in the matched data regressions exhibited a strong relationship to this outcome measure. Nonetheless, we estimated a leaner specification, including only the Hutu indicator, and it yielded even more attenuated results (coefficient of -0.58 and p-value of 0.26). Thus, the evidence of effects on attitudes toward civilian versus combatant life is quite weak statistically, despite the rather estimated odds ratios.

4.10.2 Satisfaction with the peace accords

Our next outcome variable was based on whether respondents said that they were "very satisfied" with the peace accords (=1) versus "satisfied" or "dissatisfied" (=0). The proportion responding "very satisfied" was 68% for both the matched controls and the Africare respondents. Thus, while there is a reasonable amount of variation on this question, there is no difference in rates across the two groups of respondents. The more refined estimates in Table 4.8 show as such, where the estimated coefficients are, essentially, 0. Thus, effects on attitudes toward the peace accords appear to be trivial or non-existent.

4.10.3 Support for current government and institutions

Our final attitudinal measure of political reintegration was a question about whether the government should be given more time (=1) to sort out the major problems facing Burundian society, or whether Burundians ought to consider ways to change the current government (=0). The question interacts with the political context in a number of ways that ought to be elaborated. First, the outcome of the war was one that brought about a fundamental transformation in opportunities for Hutu mobility, including the ethnic integration of major institutions such as the army, the civil service, and educational systems. At the same time, more radical Hutu parties, including FNL factions, accused the CNDD-FDD of "selling out" far short of achieving an ethnic balance in public institutions reflective of the ethnic balance in society. The ethnic quota in the military, for example, posited a 50-50 ethnic balance, far from the 85-14-1 Hutu-Tutsi-Twa balance typically presumed among Burundians. Among the ex-rebels included in our ex-combatant sample, we presume that expressions of the need to "consider ways to change the current government" would tend to reflect respondents' sympathizing with the FNL's accusations about the CNDD-FDD "selling out." In our data, we find that about 82% of the matched controls indicated that the government "should be given more time", whereas about 90% of the Africare respondents indicated as such, indicating the odds of a pro-government response are about two times higher for Africare respondents. This difference goes in the *opposite* of the hypothesized direction. The logistic regression estimates in Table 4.8 show that once we take the covariates into account, we estimate that the odds of indicating the the government should be given more time are two to four times higher, although the p-value is rather large at 0.24. Thus, if anything, the effect of the programming was to make ex-combatants more likely express impatience with the government.

4.11 Discussion

Our findings suggest that over the rather short duration that we could assess, the reintegration program produced a significant boost to income among ex-combatants who would otherwise have been among the worst off, resulting in a substantial lowering of poverty incidence. The program had a moderate effect on improving the livelihood prospects for ex-combatants, although not enough to transform substantially the overall distribution of ex-combatants over livelihood outcomes. Agricultural livelihoods dominate, and the program did not affect that. We sought to assess whether economic assistance of this form can also translate into downstream political reintegration effects, which we measured in terms of respondent attitudes toward civilian life, the peace accords, and the government. The aim was to test a key proposition underpinning current reintegration programming: that improvements to one's economic well-being may induce a more positive disposition toward political order and the laws and norms governing civilian society. We do not find compelling evidence of such downstream effects.

We recognize that there are a number of limitations to this empirical analysis. First, we are evaluating program effects within a very short timeframe. Non-Africare respondents had received their socio-economic reintegration packages no more than 9 months prior to our fieldwork. We recognize that there would be greater interest in an assessment some time later, perhaps years later. The problem is that good causal identification for such a study would be extremely unlikely. Indeed, in our case, Africare beneficiaries began to receive reintegration packages shortly after we completed fieldwork, thus erasing the discontinuity that we have exploited for our study. Second, we relied on a relatively small sample. This was due to the need to remove incidental imbalances for good causal identification. Our matching algorithm discarded over 70% of the non-Africare respondents, and this after we had already discarded Bujumbura-based respondents, who constituted 12% of the original sample. Future research ought to anticipate the need for such balance, and design more efficient sampling strategies (e.g., by sampling on already-measured community-level attributes). Third, related to the sample size issue is the fact that our analysis is confined to those registered to receive benefits outside the capital city. In many post-conflict contexts, rather large numbers of ex-combatants congregate in capital cities, and so there may be an important part of the picture that we miss here. Nonetheless, in Burundi, the vast majority of ex-combatants settled outside Bujumbura, and so we think our study is well-motivated on those grounds, not to mention the fact that excluding Bujumbura was necessary for good causal identification. Finally, we would greatly prefer to have had unobtrusive behavioral measures of political reintegration. Our assessment of political reintegration is only suggestive, and it only gets at the psychological aspect. Data on criminality and violence behavior would be ideal, if it were available. We were unable to locate fine-grained enough data for these purposes. Fourth, the discontinuity that we exploited was regional in nature. This forced us to make an assumption that, once we had controlled for individual and community-level attributes, we had achieved exchangeability all the way down to the individual level. This assumption cannot be tested.

These limitations taken into account, we feel that our study represents progress in important ways. First, the type of discontinuity that we exploit offers the *best feasible design* for measuring the impact of reintegration programs *in their totality.* Because these are "emergency" programs in fragile political environments, there is no realistic hope of a reintegration program being randomized fully or even being subject to a randomized roll-out design.[39] Our advice is that researchers seek these kinds of discontinuities, often due to program failures. We also advise researchers to seek opportunities to randomize *features* of reintegration programs—e.g. difference combinations of community-level and individual level assistance and different ways of either promoting or discouraging the creation of associations among ex-combatants.

[39]This is based on our own field experiences as well as discussions with many World Bank and United Nations program admistrators.

These kinds of within-program experiments could allow us to assess causal mechanisms and policy options. Second, while the regional nature of our discontinuity forced us to make assumptions about exchangeability across regions, it also allowed us to ensure that broader equilibrium effects were incorporated into our estimates of program effects. Income, occupation, and political attitudes are all likely to exhibit substantial local spill-over. If one had attempted to identify causal effects by comparing individuals within the same locality, these spill-overs would have been hidden. Finally, Our adjustment strategy resembles the long line of labor economics research on the United States National Supported Work program (Dehejia and Wahba, 1999; Smith and Todd, 2005). Smith and Todd (2005) have shown that program effect estimates are quite sensitive to the covariate set that one chooses. This does not seem to be such an important problem in our case, since in fact, the differences in the estimates on the raw and matched data are often small, the exception being only for one of the livelihood outcomes (no occupation). We have used a very rich covariate set that, unlike other studies, accounts for both individual characteristics and community characteristics. This is crucial, as economic and political reintegration outcomes are most certainly the result of an interaction between these two levels.

4.12 Conclusion

Ex-combatant reintegration programming is central in most transitions from civil war to peace. Considerable sums are expended on such programs worldwide. Peace researchers and program managers attribute great importance to such programs in helping societies move past critical barriers in realizing sustainable peace. These programs provide incentives that ostensibly lure ex-combatants away from the use of violence to meet their material and psychological needs, and ostensibly re-orient ex-combatants toward subscribing to norms that govern civilian society. But the evidence on the impact of these programs is very thin. There is a worrying gap between effort, expectations, and evidence.

We make a major contribution in the way of evidence. We use a bureaucratic failure in the implementation of the ex-combatant reintegration program in Burundi as a source of exogenous variation to measure the impact of the program. The nature of the failure was such that one of the implementing partners fell into dispute with the government for reasons quite independent of the characteristics those ex-combatants that were due to be served. This groups of intended beneficiaries included approximately one-third of the overall ex-combatant beneficiary population. Because of the dispute, this large group of would-be beneficiaries had their benefits withheld for nearly a year. During that time, we were able to launch a field study to measure effects of reintegration programming. We studied effects on both the proximate economic outcomes that were directly targeted ("programmatic effects"), as well as second-ordered effects on political attitudes that were expected to follow from a boost to economic well-being ("downstream effects"). To achieve good causal identification, we limited our analysis to former rebel combatants residing outside of the capital city; this subgroup actually represents the majority of Burundi's ex-combatants. We matched ex-combatants whose benefits were withheld to those that suffered no such disruption on a rich set of individual-level and community-level characteristics. We used regression adjustment to make up for the fact that the matches were not exact. We also used an inverse propensity adjustment to make up for the fact that a minority segment of ex-combatants in the region suffering the disruption had been able to access program benefits prior to the disruption.

Our assessment of "programmatic" effects is important, because past research has raised questions about whether even these proximate impacts are manifest. We find that they are. The program effected a substantial boost to the income of those who would otherwise have been among the worst off, leading to a large reduction in poverty incidence. In addition, there is moderate evidence that the livelihood and occupation outcomes were improved, although the program did not effect any sort of transformation in the overall distribution of ex-combatants over occupations. We do not find compelling evidence for "downstream" effects of such economic assistance on political reintegration, at least as measured through attitudes. While we found a moderate effect on ex-combatants' sense that civilian life was preferable to combatant life, there was no effect of either increasing levels of satisfaction with the peace accords or increasing levels of support for current governing institutions. Thus, literature suggesting that a boost to economic well-being can induce a more positive orientation toward stability and norms governing civilian society may be too hopeful. Most probably, more direct interventions, via media or counseling, are probably required to shift these attitudes.

We hope that our research sets provides an example for research designs that successfully exploit unusual programmatic discontinuities to estimate important causal effects. Sometimes these discontinuities happen by design, as when programs are rolled out in a random fashion. Sometimes the discontinuities happen by accident, as in the program we study in this paper. We suspect that bureaucratic failures, halts to service delivery, and other such unanticipated sources of variation in program performance exist all over the place. These provide excellent opportunities to measure program effects, particularly if the disruptions are ones that last for a long time. Such studies should be designed to ensure adequate power, and should attempt to incorporate unobtrusive behavioral measures to improve the tangibility of findings.

References

Abadie A, Imbens GW. 2008. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica.* 76(6):1537-1557.

Alden C. 2002. "Making old soldiers fade away: Lessons from the reintegration of demobilized soldiers in Mozambique." *Security Dialogue.* 33(3):341-356.

Angrist JD, Pischke JS. 2008. *Mostly Harmless Econometrics.* Princeton: Princeton University Press.

Annan J, Patel AC. 2009. *Critical Issues and Lessons in Social Reintegration: Balancing Justice, Psychological Well Being, and Community Reconciliation.* Paper prepared for the First International Congress on Disarmament, Demobilization, and Reintegration, Cartagena.

Baare A. 2006. *An Analysis of Transitional Economic Reintegration.* Swedish Initiative for Disarmament, Demobilisation and Reintegration (SIDDR). Ministry of Foreign Affairs, Sweden.

Blattman C, Annan J. 2009. "The consequences of child soldiering." *Review of Economics and Statistics.* Forthcoming.

Boshoff H, Vrey W. 2006. *Disarmament, Demobilisation and Reintegration During the Transition in Burundi: A Technical Analysis.* Pretoria: Institute for Security Studies, Monograph Number 125.

Boutros-Ghali B. 1995. *An Agenda for Peace.* New York, NY: United Nations

Bryden A, Hanggi H, eds. 2005. *Security Governance in Post-Conflict Peacebuilding.* New Brunswick, NY: Transaction Publishers.

Castillo G. 2008. *Rebuilding War-Torn States: The Challenges of Post-Conflict Reconstruction.* New York, NY. Oxford University Press.

Colletta N, Kostner M, Weidehofer I. 1996. *The Transition from War to Peace in Sub-Sahara Africa*. Washington, DC: World Bank.

Collier P, Hoeffler A. 2004. "Greed and grievance in civil war." *Oxford Economic Papers*. 56(4):563-595.

Collier P. 1994. "Demobilization and insecurity: A study in the economics of the transition from war to peace." *Journal of International Development*. 6(3):343-351.

Cousens E, Kumar C, Wermester K. 2001. *Peacebuilding as Politics: Cultivating Peace in Fragile Societies*. Boulder, CO: Lynne Rienner.

Dehejia R, Wahba S. 1999. "Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs." *Journal of the American Statistical Association*. 94(448):1053-1062.

Diamond A, Sekhon JS. 2005. "Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies." Typescript, Harvard University and University of California, Berkeley.

DiNardo J, Lee DS. 2010. "Program evaluation and research designs." *NBER Working Paper*. Number 16016.

Douma P, Gasana JM, Specker L. 2008. *Reintegration in Burundi: Between Happy Cows and Lost Investments*. Typescript, Conflict Resolution Unit of the Netherlands Institute of International Relations, "Clingendael."

Doyle M, Sambanis N. 2006. *Making War and Building Peace: United Nations Peace Operations*. Princeton, NJ: Princeton University Press

Feil S. 2004. "Laying the foundations: Enhancing security capabilities." In Orr RC, ed. *Winning the Peace: An American Strategy for Post-Conflict Reconstruction*. Washington, DC: The Center for Strategic and International Studies.

Fortna P. 2008. *Does Peacekeeping Work? Shaping Belligerents' Choices after Civil War.* Princeton, NJ: Princeton University Press

Geenen S. 2007. "Former combatants at the crossing: How to assess the reintegratoin of former combatants in the security and development nexus? Case study: Ruyigi (Burundi) and Kinshasa (DRC)." Typescript, University of Antwerp.

Green DP, Ha SE, Bullock JG. 2010. "Enough already about 'black box' experiments: Studying mediation is more difficult than most scholars suppose." *The Annals of the American Academy of Political and Social Science.* 628(1):200-208.

Hanson S. 2007. *Disarmament, Demobilization, and Reintegration (DDR) in Africa.* Backgrounder. http://www.cfr.org/publication/12650/

Ho D, et al. 2007. "Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference." *Political Analysis.* 15:199-236.

Ho D, et al. 2009. "MatchIt: Nonparametric preprocessing for parametric causal inference." Typescript, Stanford University and Harvard University.

Humphreys M, Weinstein J. 2007. "Demobilization and reintegration." *Journal of Conflict Resolution.* 51(4):531-567.

Jennings K. 2007. "The struggle to satisfy: DDR through the eyes of ex-Combatants in Liberia." *International Peacekeeping.* 14(2):1-15.

Jennings K. 2008. "Unclear ends, unclear means: Reintegration in postwar societies—the case of Liberia." *Global Governance.* 14:327-345.

King G, Zeng L. 2006. "The dangers of extreme counterfactuals." *Political Analysis.* 14(2):131-159.

King G, et al. 2001. "Analyzing incomplete political science data: An alternative algorithm for multiple imputation." *American Political Science Review.* 95(1):49-69.

Kingma K, Pauwels N. 2000. "Demobilization and reintegration in the 'downsizing decade."' In Pauwels N, ed. *War Force to Work Force: Global Perspectives on Demobilization and Reintegration.* Baden-Baden: Nomos.

Kingma K. 1997. "Demobilization of combatants after civil wars in Africa and their reintegration into civilian life." *Policy Sciences* 30(3):151-166.

Koenker R, Hallock KF. 2000. "Quantile regression: an introduction. [Long version.]" Typescript, University of Illinois at Urbana-Champaign.

Korn EL, Graubard BI. 1999. *Analysis of Health Surveys.* New York: Wiley.

Kumar K, ed. 1997. *Rebuilding Societies After Civil War: Critical Roles for International Assistance.* Boulder, CO: Lynne Rienner.

Levely I. 2010. "Disarmament, demobilization, reinsertion and reintegration in Post-War Liberia: A failed approach or simply a failed program?" Typescript, Charles University, Prague.

Little RA, Rubin DB. 2002. *Statistical Analysis with Missing Data, 2nd Edition.* New York: Wiley.

Long JS. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks: Sage.

McMullin J. 2004. "Reintegration of combatants: were the right lessons learned in Mozambique?" *International Peacekeeping.* 11(4): 625-643.

Muggah R. 2005. "No magic bullet: A critical perspective on disarmament, demobilization and reintegration (DDR) and weapons reduction in post-conflict contexts." *The Round Table.* 379:239-252.

Muggah R. 2009. *Security and Post-Conflict Reconstruction: Dealing with Fighters in the Aftermanth of War.* New York, NY: Routledge.

Multi-country Demobilization and Reintegration Program [MDRP]. 2007. *MDRP Quarterly Progress Report, April-June 2007.* Washington, DC: MDRP Secretariat.

Murphy KM, Topel RH. 1985. "Estimation and inference in two-step econometric models." *Journal of Business and Economic Statistics.* 3:370-379.

National Commission for Disarmament, Demobilization and Reintegration [CNDRR]. 2005. *Ganuka: Lettre d'information.* No. 1. January, 2005.

Nindorera W. 2007. "Security sector reform in Burundi: Issues and challenges for improving civilian protection." *Centre dAlerte et de Prévention des Conflits and North-South Institute Working Paper.* Bujumbura: CENAP. Available at www.nsi-ins.ca

Paes WC. 2005. "The Challenges of disarmament, demobilization and reintegration in Liberia." *International Peacekeeping.* 12(2): 253-61.

Paris R. 2004. *At War's End: Building Peace After Civil Conflict.* New York, NY: Cambridge University Press.

Pouligny B. 2004. *The Politics and Anti-Politics of Contemporary Disarmament, Demobilization and Reintegration Programs.* Paris: CERI/SGDN.

Pugel J. 2009. "Measuring reintegration in Liberia: assessing the gap between outputs and outcomes." In Muggah R (2009).

Reno W. 1999. *Warlord Politics and African States.* Boulder, CO: Lynne Rienner.

Republic of Burundi. 2004. *Joint Operations Plan for Pre-Disarmament, Disarmament, Combatant Verification, and Demobilization.* [Official document, Bujumbura, November 9, 2004.]

Rizopoulos D, Moustaki I. 2008. "Generalized latent variable models with non-linear effects." *British Journal of Mathematical and Statistical Psychology*. 61(2):415-438.

Rizopoulos D. 2006. "ltm: An R package for latent variable modeling and item response theory analyses." *Journal of Statistical Software*. 17(5).

Rubin DB. 1976. "Multivariate matching methods that are equal percent bias reducing, II: Maximums on bias reduction for fixed sample sizes." *Biometrics*. 32(1):121-132.

Rubin DB. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Rubin DB, Thomas N. 2000. "Combining propensity score matching with additional adjustments for prognostic covariates." *Journal of the American Statistical Association*. 95(450):573-585.

Rubin B. 2003. *Identifying Options and Entry Points for Disarmament, Demobilization, and Reintegration in Afghanistan*. New York: Center on International Cooperation.

Samii C. 2010. "Missing data." To appear in Badie B, et al. *International Encyclopedia of Political Science*. New York: Sage.

Sekhon JS. 2009. "Opiates for the matches: Matching methods for causal inference." *Annual Review of Political Science*. 12:487-508.

Sekhon JS. n.d. "Multivariate and propensity score matching software with automated balance optimization: The Matching package for R." *Journal of Statistical Software*. (Forthcoming.)

Smith JA, Todd PE. 2005. "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of Econometrics*. 125(1-2):305-353.

Soderstrom J. 2010. *Reintegrating Ex-combatants in Liberia - An Opportunity for Democracy?* Unpublished thesis. Uppsala University.

Spear J. 2002. "Disarmament and demobilization." In Stedman et al, 2005.

Spear J. 2006. "From political economies of war to political economies of peace: The contribution of DDR after wars of predation." *Contemporary Security Policy.* 27(1):168-189.

Specker L. 2008. *The R-Phase of DDR processes.* Netherlands Institute of International Relations 'Clingendael' Conflict Research Unit.

Stedman SJ, Rothchild D, Cousens EM, eds. 2002. *Ending Civil Wars: The Implementation of Peace Agreements.* Boulder, CO: Lynne Rienner.

Tajima Y. 2009. *Background Paper on Economic Reintegration.* Prepared for the First International Congress on Disarmament, Demobilization, and Reintegration, Cartagena.

Tomz M, et al. 2003. "Clarify: Software for interpreting and presenting statistical results." *Journal of Statistical Software.* 8(1).

United Nations Development Program (UNDP). 2000. *Sharing Ground in Post-Conflict Situations: The Role of UNDP in Support of Reintegration Programmes.* New York:UNDP.

United Nations. 2000. *The Role of United Nations Peacekeeping in Disarmament, Demobilization and Reintegration. Report of the Secretary-General to the Security Council, S/2000/101.* New York: United Nations.

United Nations. 2006. *Integrated Disarmament Demobilization and Reintegration Strategies (IDDRS) 2.10: The UN Approach to DDR.* New York: United Nations.

Uvin P. 2007. *Ex-combatants in Burundi: Why They Joined, Why They Left, How They Fared.* Washington, DC: Multi-country Demobilization and Reintegration Program/World Bank.

van Buuren S, Groothuis-Oudshoorn K. 2010. "MICE: Multivariate imputation by chained equations in R." *Journal of Statistical Software.* Forthcoming.

van der Ark LA. 2007. "Mokken scale analysis in R." *Journal of Statistical Software.* 20(11).

Walter B. 2001. *Committing to Peace: The Successful Settlement of Civil Wars.* Princeton, NJ: Princeton University Press

World Bank. 2002. *Greater Great Lakes Regional Strategy for Demobilization and Reintegration.* Report No. 238669-AFR. Washington, DC: World Bank and MDRP Secretariat.

World Bank. 2003. *Breaking the Conflict Trap.* Washington, DC: World Bank.

World Bank. 2004a. *Position Paper: Targeting MDRP Assistance: Ex-Combatants and Other War-Affected Populations.* Washington, DC: World Bank.

World Bank. 2004b. *Technical Annex for a Proposed Grant of SDR 22.2 Million to the Republic of Burundi for an Emergency Demobilization, Reinsertion and Reintegration Program.* Washington, DC: World Bank.

World Bank. 2009. *Emergency Project Paper on a Proposed Emergency Recovery Grant in the Amount of SDR 10,1 Million to the Republic of Burundi for an Emergency Demobilization and Transitional Reintegration Project.* Washington, DC: World Bank.

Tables

Table 4.1: **Estimated Sizes of Armed Forces as of January 2004**

Name of Force	Estimated Size
Forces Armées Burundaises (national army)	45,000
CNDD-FDD I (Nkurunziza faction)	25,000
CNDD-FDD II (Ndayikengurukiye faction)	3,000
FNL-PALIPEHUTU I (Rwasa faction)*	3,000
CNDD (Nyangoma faction)	1,000
FNL-PALIPEHUTU II (Mugabarabona faction)	1,000
FROLINA (Kalumba faction)	1,000
PALIPEHUTU (Karatasi faction)	1,000
Total	80,000

Source: World Bank (2004), p. 17.

**Not a party to the peace process until September 2008.*

Table 4.2: **Military Status of Surviving Civil War Combatants as of June 2007**

Status	Number	Of which from rebel forces
In Forces de Defense Nationale (new national army)	28,390	approx. 9,000
In Police Nationale	approx. 20,000	approx. 8,000
Demobilized	23,185	approx. 14,000*
In FNL-PALIPEHUTU (active factions)	approx. 6,500**	approx. 6,500**
Totals	approx. 78,075	approx. 37,500

Sources: MDRP (2007), Nindorera (2007), World Bank (2009).

* *Approximately 9,000 out of an initial 14,000 demobilized in a first phase (prior to 2004) were from the rebel groups. Of the remaining approximately 9,000 that were demobilized (to generate the 23,185 total), personal communications from program staff suggest that about 5,000 of them were from the rebel groups, although this has not been documented to our knowledge.*

** *This figure is based on the demobilization targets in World Bank (2009), which were estimated some time after June 2007. There has been some controversy as to whether this overstated the number of "true" FNL-PALIPEHUTU fighters that remained into 2008-9 by a few thousand. Without any other evidence, this is our best estimate, but it may be biased slightly upward.*

Table 4.3: **Program Access in Africare and Non-Africare Regions, Before and After NGO transition in Fall 2006 and as of the Time of Fieldwork in July 2007**

Region	Pre-transition cases completed by 12/06	Remaining caseload...	...% subject to disruption	Overall disruption rate
Africare provinces	1,982	2,257	100%	.53
Non-Africare provinces	3,213	5,925	0%	0

Sources: World Bank and Multi-country Demobilization and Reintegration Program (MDRP), Quarterly Reports, October-December 2006 and April-June 2007, and Africare Annual Report, 2008. Africare provinces include Cankuzo, Gitega, Karuzi, Muramvya, Mwaro, and Ruyigi. The table refers only to ex-combatants who are (1) male, constituting 97.5% of ex-combatants in the reintegration program, and (2) registered to receive program benefits outside Bujumbura.

Table 4.4: **Balance Statistics**

		Before matching					After matching					% Improvement	
	Variable	**Means treated**	**Means control**	**Var. Ratio**	**t-test p-value**	**KS test p-value**	**Means treated**	**Means control**	**Var. Ratio**	**t-test p-value**	**KS test p-value**	**Mean Diff.**	**eQQ Mean**
Imputation 1	Age	34.37	33.31	1.20	0.34	0.20	34.37	33.20	1.58	0.39	0.27	-10.39	-60.92
Sample sizes:	Hutu	0.95	0.96	1.08	0.89		0.95	0.96	1.23	0.77		-174.74	0.00
Africare	Father in agriculture, prewar	0.82	0.69	0.70	0.01		0.82	0.82	0.99	1.00		100.00	53.09
n=110	Prewar education	4.75	5.10	0.85	0.24	0.42	4.75	4.93	1.24	0.62	0.60	48.10	-37.45
Non-Africare	Prewar wealth	-0.05	0.10	1.18	0.03	0.00	-0.05	-0.02	1.07	0.80	0.81	84.37	62.11
Before matching	Non-comm. officer	0.72	0.66	0.91	0.26		0.72	0.72	0.99	1.00		100.00	76.55
n = 261	Comm. officer	0.11	0.18	0.65	0.05		0.11	0.14	0.82	0.60		63.55	17.91
After matching	Unit death rate	0.09	0.13	0.59	0.00	0.11	0.09	0.08	1.46	0.48	0.79	73.18	67.38
n = 67	Years in faction	9.60	10.71	0.62	0.00	0.00	9.60	9.88	1.81	0.46	0.07	74.50	20.56
	Family death rate	0.15	0.18	1.05	0.13	0.17	0.15	0.13	1.76	0.46	0.17	43.58	27.42
	Non-CNDD-FDD	0.09	0.22	0.48	0.00		0.09	0.09	0.99	1.00		100.00	64.82
	Demob. Date (std.)	-0.05	0.04	0.08	0.27	0.87	-0.05	-0.08	1.13	0.57	0.51	67.83	67.34
	War violence index	5.17	9.31	0.32	0.00	0.00	5.17	5.16	1.96	0.99	0.00	99.88	54.07
	Log(pop. density)	5.63	5.76	1.20	0.01	0.00	5.63	5.72	1.21	0.14	0.03	30.97	12.49
	Ruling party province	0.65	0.59	0.94	0.21		0.65	0.65	0.98	0.90		86.70	-2.61
	Propensity score	0.40	0.25	1.53	0.00	0.00	0.40	0.38	1.54	0.50	0.11	88.75	57.29
Imputation 2	Age	34.37	33.31	1.20	0.34	0.19	34.37	33.29	1.30	0.45	0.26	-1.83	-52.22
Sample sizes:	Hutu	0.95	0.96	1.08	0.89		0.95	0.96	1.23	0.77		-174.74	0.00
Africare	Father in agriculture, prewar	0.81	0.68	0.72	0.01		0.81	0.82	1.03	0.88		92.85	64.82
n=110	Prewar education	4.75	5.10	0.85	0.24	0.37	4.75	4.63	1.00	0.76	0.98	66.27	46.55
Non-Africare	Prewar wealth index	-0.06	0.10	1.19	0.02	0.00	-0.06	0.04	1.04	0.31	0.07	38.31	1.38
Before matching	Non-comm. officer	0.72	0.66	0.91	0.26		0.72	0.74	1.04	0.79		69.28	53.09
n = 261	Comm. officer	0.11	0.18	0.65	0.05		0.11	0.11	0.99	1.00		100.00	79.48
After matching	Unit death rate	0.09	0.12	0.57	0.01	0.17	0.09	0.07	1.62	0.37	0.66	61.40	54.24
n = 67	Years in faction	9.60	10.68	0.61	0.00	0.00	9.60	9.95	1.83	0.36	0.13	67.12	19.28
	Family death rate	0.15	0.18	1.05	0.13	0.13	0.15	0.15	1.30	0.89	0.12	88.49	-23.30
	Non-CNDD-FDD	0.09	0.22	0.48	0.00		0.09	0.09	0.99	1.00		100.00	53.09
	Demob. Date (std.)	-0.05	0.04	0.08	0.26	0.74	-0.05	-0.07	2.76	0.55	0.30	73.87	66.21
	War violence index	5.17	9.31	0.32	0.00	0.00	5.17	4.97	2.09	0.77	0.00	95.21	56.54
	Log(pop. density)	5.63	5.76	1.20	0.01	0.00	5.63	5.69	1.19	0.31	0.05	51.46	27.62
	Ruling party province	0.65	0.59	0.94	0.21		0.65	0.65	0.99	1.00		100.00	38.43
	Propensity score	0.40	0.25	1.47	0.00	0.00	0.40	0.38	1.46	0.40	0.04	85.63	50.29
Imputation 3	Age	34.37	33.31	1.20	0.34	0.19	34.37	32.81	1.93	0.22	0.16	-47.18	-93.92
Sample sizes:	Hutu	0.95	0.96	1.08	0.89		0.95	0.97	1.63	0.52		-449.47	NA
Africare	Father in agriculture, prewar	0.89	0.69	0.69	0.01		0.82	0.84	1.08	0.75		86.26	69.01
n=110	Prewar education	4.75	5.10	0.85	0.24	0.42	4.75	5.16	1.42	0.23	0.34	-19.37	-4.49
Non-Africare	Prewar wealth index	-0.05	0.10	1.18	0.03	0.01	-0.05	-0.07	1.12	0.77	0.71	81.87	51.75
Before matching	Non-comm. officer	0.72	0.66	0.91	0.26		0.72	0.75	1.09	0.59		38.55	77.87
n = 261	Comm. officer	0.11	0.18	0.65	0.05		0.11	0.11	0.99	1.00		100.00	80.63
After matching	Unit death rate	0.09	0.12	0.61	0.03	0.28	0.09	0.07	1.55	0.19	0.39	33.29	35.34
n = 71	Years in faction	9.60	10.78	0.56	0.00	0.00	9.60	9.90	1.40	0.46	0.03	74.53	33.44
	Family death rate	0.15	0.18	1.05	0.13	0.14	0.15	0.12	1.35	0.40	0.67	32.57	30.06
	Non-CNDD-FDD	0.09	0.22	0.48	0.00		0.09	0.09	0.99	1.00		100.00	88.93
	Demob. Date (std.)	-0.05	0.04	0.08	0.27	0.65	-0.05	-0.08	2.27	0.45	0.19	65.32	58.23
	War violence index	5.17	9.31	0.32	0.00	0.00	5.17	5.29	1.90	0.86	0.00	96.93	52.36
	Log(pop. density)	5.63	5.76	1.19	0.01	0.00	5.63	5.71	1.24	0.16	0.01	33.84	9.91
	Ruling party province	0.65	0.59	0.94	0.21		0.65	0.65	0.99	1.00		100.00	80.63
	Propensity score	0.39	0.26	1.46	0.00	0.00	0.39	0.38	1.56	0.50	0.07	88.69	58.77
Imputation 4	Age	34.37	33.31	1.20	0.34	0.15	34.37	33.30	1.41	0.44	0.35	-0.97	-41.41
Sample sizes:	Hutu	0.95	0.96	1.08	0.89		0.95	0.96	1.23	0.76		-174.74	0.00
Africare	Father in agriculture, prewar	0.81	0.69	0.72	0.01		0.81	0.82	1.03	0.88		92.63	65.34
n=110	Prewar education	4.75	5.10	0.85	0.24	0.36	4.75	4.70	1.25	0.90	0.37	87.03	5.95
Non-Africare	Prewar wealth index	-0.06	0.10	1.18	0.03	0.00	-0.06	-0.05	1.00	0.99	0.68	99.21	53.79
Before matching	Non-comm. officer	0.72	0.66	0.91	0.26		0.72	0.71	0.98	0.90		84.64	53.78
n = 261	Comm. officer	0.11	0.18	0.65	0.05		0.11	0.11	0.99	1.00		100.00	79.78
After matching	Unit death rate	0.10	0.12	0.68	0.07	0.28	0.10	0.08	1.39	0.52	0.99	58.52	74.28
n = 68	Years in faction	9.60	10.77	0.59	0.00	0.00	9.60	10.05	1.89	0.24	0.05	61.93	12.22
	Family death rate	0.15	0.18	1.05	0.13	0.14	0.15	0.16	1.52	0.54	0.05	51.23	-58.19
	Non-CNDD-FDD	0.09	0.22	0.48	0.00		0.09	0.09	0.99	1.00		100.00	65.34
	Demob. Date (std.)	-0.05	0.04	0.08	0.27	0.75	-0.05	-0.07	1.83	0.67	0.42	79.32	76.21
	War violence index	5.17	9.31	0.32	0.00	0.00	5.17	4.88	2.03	0.68	0.01	93.11	61.81
	Log(pop. density)	5.63	5.75	1.18	0.01	0.00	5.63	5.70	1.23	0.23	0.05	42.77	29.23
	Ruling party province	0.65	0.59	0.94	0.21		0.65	0.64	0.97	0.81		73.39	-21.32
	Propensity score	0.39	0.26	1.47	0.00	0.00	0.39	0.38	1.60	0.47	0.15	87.86	62.25
Imputation 5	Age	34.37	33.31	1.20	0.34	0.19	34.37	32.87	1.50	0.28	0.22	-41.19	-71.79
Sample sizes:	Hutu	0.95	0.96	1.08	0.89		0.95	0.95	0.99	1.00		100.00	0.00
Africare	Father in agriculture, prewar	0.81	0.69	0.72	0.01		0.81	0.81	0.99	1.00		100.00	64.82
n=110	Prewar education	4.75	5.10	0.85	0.24	0.36	4.75	4.82	1.24	0.84	0.85	79.24	23.64
Non-Africare	Prewar wealth index	-0.06	0.09	1.19	0.03	0.00	-0.06	-0.11	1.06	0.55	0.83	62.75	66.83
Before matching	Non-comm. officer	0.72	0.66	0.91	0.26		0.72	0.73	1.01	0.90		84.64	76.55
n = 261	Comm. officer	0.11	0.18	0.65	0.05		0.11	0.11	0.99	1.00		100.00	79.48
After matching	Unit death rate	0.08	0.12	0.61	0.01	0.08	0.08	0.07	2.19	0.20	0.80	47.94	45.50
n = 67	Years in faction	9.60	10.74	0.62	0.00	0.00	9.60	10.13	1.46	0.19	0.01	53.57	4.34
	Family death rate	0.15	0.18	1.05	0.13	0.15	0.15	0.15	1.29	0.92	0.25	91.76	5.90
	Non-CNDD-FDD	0.09	0.22	0.48	0.00		0.09	0.10	0.91	0.84		93.08	64.82
	Demob. Date (std.)	-0.05	0.04	0.08	0.28	0.71	-0.05	-0.10	1.83	0.26	0.32	43.94	70.11
	War violence index	5.17	9.31	0.32	0.00	0.00	5.17	5.38	1.78	0.77	0.00	94.90	47.88
	Log(pop. density)	5.63	5.75	1.17	0.01	0.00	5.63	5.72	1.16	0.15	0.04	27.35	25.44
	Ruling party province	0.65	0.59	0.94	0.21		0.65	0.65	0.99	1.00		100.00	58.96
	Propensity score	0.40	0.25	1.52	0.00	0.00	0.40	0.38	1.46	0.42	0.04	86.23	55.76

Table 4.5: **Estimates from OLS regressions on log(income/month + 1)**

	No adjustment			Matching & regression			Matching, regression, & het. exposure adjustment		
	coef.	s.e.	*p* -val.	coef.	s.e.	*p* -val.	coef.	s.e.	*p* -val.
Africare	-0.67	0.33	0.04	-0.59	0.53	0.26	-1.12	1.00	0.26
Age				-0.06	0.04	0.14	-0.06	0.04	0.14
Hutu				3.31	2.10	0.12	3.31	2.10	0.12
Father in agriculture, prewar				-1.18	0.74	0.11	-1.18	0.74	0.11
Prewar education				0.17	0.13	0.22	0.17	0.13	0.22
Prewar wealth index				0.52	0.50	0.30	0.52	0.50	0.30
Non-comm. officer				-0.19	0.51	0.71	-0.19	0.51	0.71
Comm. officer				0.07	0.88	0.93	0.07	0.88	0.93
Unit death rate				6.30	3.40	0.07	6.30	3.40	0.07
Years in faction				0.09	0.17	0.62	0.09	0.17	0.62
Family death rate				1.84	1.25	0.14	1.84	1.25	0.14
Non-CNDD-FDD				0.60	1.07	0.58	0.60	1.07	0.58
Demob. Date (std.)				-0.83	0.63	0.19	-0.83	0.63	0.19
War violence index				0.16	0.09	0.09	0.16	0.09	0.09
Log(pop. density)				-0.14	0.68	0.84	-0.14	0.68	0.84
Ruling party province				-1.19	0.59	0.04	-1.19	0.59	0.04
Propensity score				8.67	5.34	0.11	8.67	5.34	0.11
Constant	9.37	0.20	0.00	4.15	6.61	0.53	4.15	6.61	0.53
N from imputed datasets	371, 371, 371, 371, 371			177, 177, 181, 178, 177			177, 177, 181, 178, 177		

Standard errors computed with clustering at commune level.

Table 4.6: **Estimates from quantile regressions on log(income/month+1)**

	No adjustment			Matching & regression			Matching, regression, & het. exposure adjustment		
	coef.	s.e.	p-val.	coef.	s.e.	p-val.	coef.	s.e.	p-val.
Decile 1	-0.92	2.89	0.75	-1.01	2.24	0.65	-1.90	4.22	0.65
Decile 2	-0.69	0.24	0.00	-0.54	0.98	0.59	-1.01	1.85	0.59
Decile 3	-0.69	0.16	0.00	-0.60	0.31	0.06	-1.14	0.58	0.06
Decile 4	-0.41	0.17	0.02	-0.51	0.28	0.07	-0.97	0.52	0.07
Decile 5	-0.64	0.17	0.00	-0.44	0.24	0.07	-0.83	0.45	0.07
Decile 6	-0.60	0.21	0.01	-0.43	0.18	0.02	-0.80	0.34	0.02
Decile 7	-0.41	0.14	0.00	-0.39	0.18	0.03	-0.74	0.33	0.03
Decile 8	-0.51	0.18	0.00	-0.37	0.22	0.11	-0.70	0.42	0.11
Decile 9	-0.25	0.17	0.14	-0.35	0.21	0.10	-0.66	0.40	0.10
N from imputed datasets	371,371,371,371,371			177,177,181,178,177			177,177,181,178,177		

Standard errors computed using robust inverted rank test intervals. Regressions on matched data include the covariates in Table 4.5. The coefficients on these covariates are not displayed to save space.

Table 4.7: **Estimates from multinomial and ordered logistic regressions on occupation categories**

	No adjustment				Matching & regression				Matching, regression, & het. exposure adjustment			
Unordered outcomes (multinomial logit)	coef.	RRR	s.e.	*p*-val.	coef.	RRR	s.e.	*p*-val.	coef.	RRR	s.e.	*p*-val.
No occ. vs. agr.	-0.57	0.56	0.43	0.18	1.02	2.78	1.22	0.42	1.93	6.86	2.30	0.42
Skilled occ. vs. agr.	-1.00	0.37	0.46	0.03	-0.94	0.39	0.78	0.23	-1.78	0.17	1.46	0.23
Unskilled occ. vs. agr.	-0.33	0.72	0.47	0.48	-0.52	0.59	0.56	0.35	-0.99	0.37	1.06	0.35
Ordered outcomes (ordered logit)	coef.	OR	s.e.	*p*-val.	coef.	OR	s.e.	*p*-val.	coef.	OR	s.e.	*p*-val.
1=no job, 2=agriculture or unskilled, 3=skilled	-0.23	0.80	0.20	0.25	-0.77	0.46	0.50	0.13	-1.46	0.23	0.95	0.13
N from imputed datasets	371, 371, 371, 371, 371				177, 177, 177, 181, 178, 177				177, 177, 177, 181, 178, 177			

Multinomial logit standard errors are not clustered at commune level because of constraints in our imputation combination software. But we estimated the models on each of the imputation-completed datasets with and without clustered standard errors, and then compared results. They were nearly identical and no inferences changed. Regressions on matched data include the covariates in Table 4.5. The coefficients on these covariates are not displayed to save space. "RRR" refers to relative risk ratio and "OR" to odds ratio.

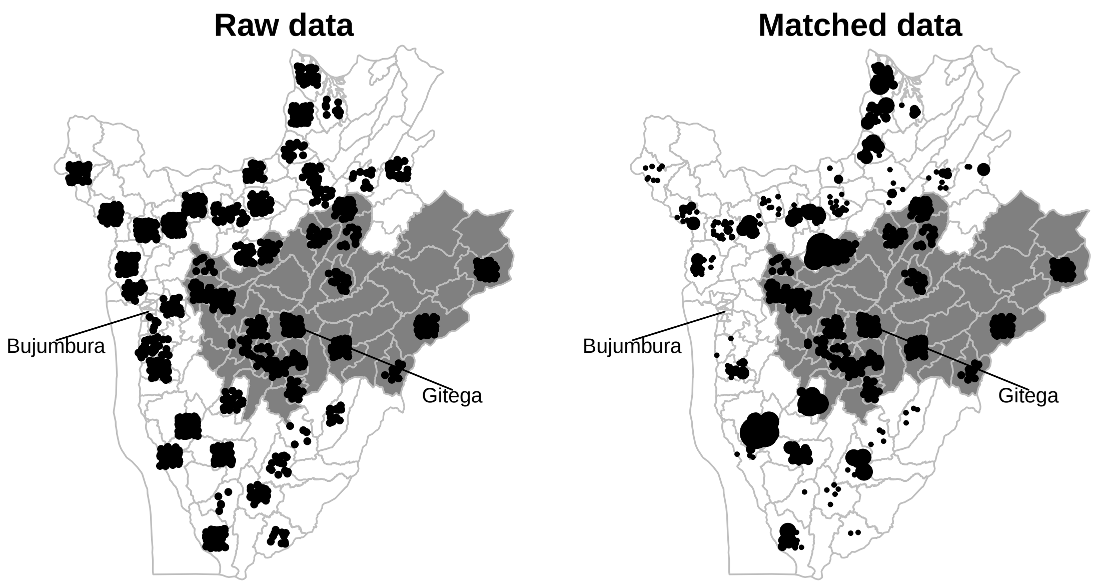
Table 4.8: **Estimates from logistic regressions on political attitudes outcomes**

	No adjustment				Matching & regression				Matching, regression, & het. exposure adjustment			
Outcome	coef.	OR	s.e.	*p* -val.	coef.	OR	s.e.	*p* -val.	coef.	OR	s.e.	*p* -val.
i. "Life better as civilian than combatant"	-0.06	0.94	0.29	0.83	-0.72	0.49	0.51	0.17	-1.36	0.26	0.96	0.17
ii. "Very satisfied with peace accords"	0.05	1.05	0.23	0.83	-0.02	0.99	0.39	0.97	-0.03	0.97	0.74	0.97
iii. "Government should have more time to solve problems"	0.19	1.21	0.35	0.60	0.74	2.11	0.68	0.29	1.41	4.08	1.29	0.29
N from imputed datasets	371, 371, 371, 371, 371				177, 177, 181, 178, 177				177, 177, 181, 178, 177			

Standard errors computed with clustering at commune level. Regressions on matched data include the covariates in Table 4.5. The coefficients on these covariates are not displayed to save space. "OR" refers to odds ratio.

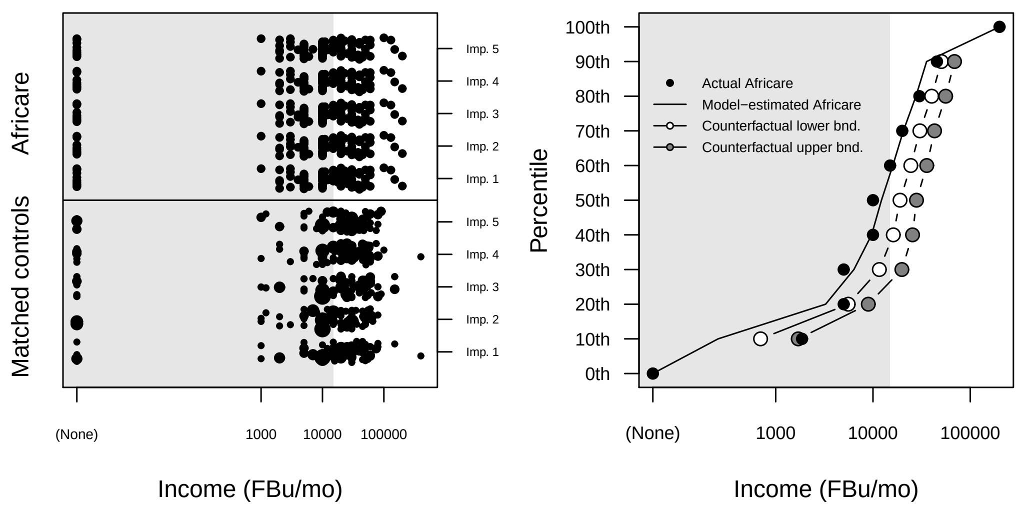
Figures

Figure 4.1: **Subject locations**



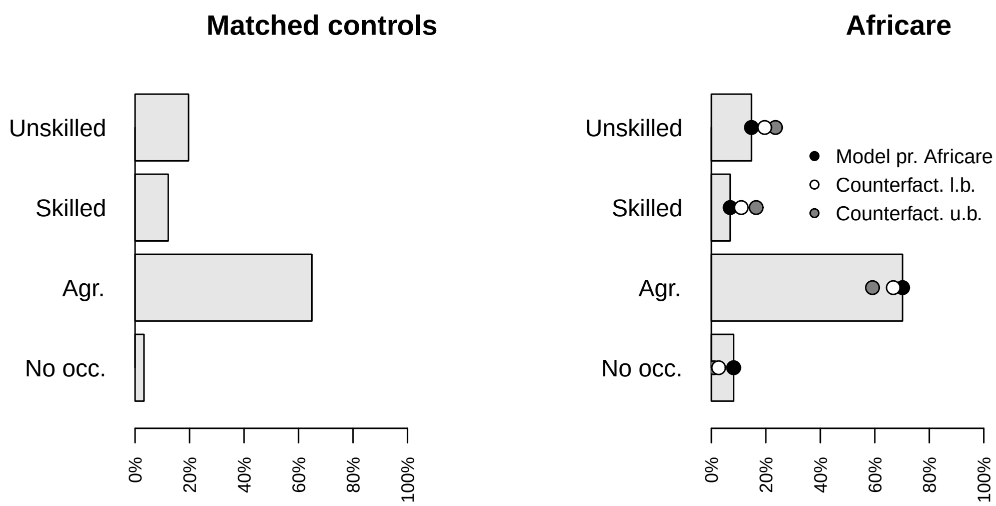
The gray region in the maps is Africare's domain of responsibility for the reintegration program. The boundaries shown on the maps are commune boundaries. The left map shows the locations of subjects in the raw data. The right map shows the locations of subjects in the matched data. The radii of the non-Africare points are proportional to the weight these subjects receive in the analysis. Observations from all five of the imputation-completed datasets are shown together in the maps.

Figure 4.2: **Differences in log-income distributions and estimated effects on the cumulative log-income distribution**



The figure on the left shows the distribution of income (on the natural log scale) for Africare and matched control (non-Africare) observations. Distributions are shown for each of the 5 imputation-complete datasets. The size of the points for the matched controls is proportional to the weight assigned to that observation after matching. The figure on the right shows the cumulative log-income distribution for Africare respondents, and then predicted counterfactual distributions. The gray zone corresponds to points below the 15000FBu/month ($1.25/day at PPP) poverty line.

Figure 4.3: **Livelihood outcome distributions and estimated effects**



The left graph shows the livelihood outcome distribution for the matched controls. The right graph shows estimated counterfactual outcomes imposed over the outcome distribution for the Africare-region ex-combatants. The estimated effect is the difference between the modeled "actual" outcomes (the solid black dots, which show in-sample fitted values) and the estimated counter-factual outcomes. The white dots show the lower-bound counter-factual estimates, and the gray dots shows the upper bound counter-factual estimates.