Repressed Productive Potential and Revolt: Insights from an Insurgency in Burundi

Appendix of Supporting Information

Contents

A Formal model

We consider how repression and horizontal inequality interact with individual economic conditions to induce preferences to revolt. Our analysis builds on a "tipping model" of insurgent participation (Medina, 2007, 2013; Schelling, 1978; Wood, 2003). Suppose a society under the domination of an incumbent ruling group. Members of a non-ruling group are indexed by i. Each non-ruling group member is characterized by a value, $a_i > 0$, capturing lifetime, unfettered productive potential. Status quo policies, enforced by the ruling group, are conceptualized in terms of a parameter, $0 < \tau < 1$, that defines the share of the non-ruling group members' product that they get to keep for themselves, with the remainder hoarded by the ruling group. Barriers to mobility under the status quo become targets of political change for members of repressed groups, and such desired change may provide motivation for revolt (DeNardo, 1985; Glaeser, Ponzetto and Shleifer, 2007; Medina, 2007; Wood, 2003). The non-ruling group members face a choice of whether to participate (P) in a revolt or not (N) , which we denote in terms of a strategy $s_i \in \{P, N\}$, respectively. The probability that a revolt will be successful is given by $\omega(f_t)$, where f_t is the share of the non-ruling group that is participating as of period t. Participation comes at an individual-specific cost, $c_i > 0$. This term captures net costs of joining, regardless of the outcome of the revolt. One could think of this as a reduced form way of expressing opportunities forgone, assuming that under the condition of not joining, the individual is choosing the next best option (which could even include leaving the country).

If the revolt fails, the regime does not change. Those who participated in the failed revolt pay the cost, c_i, and receive no benefit. Those who did not participate in the failed revolt continue to receive the status quo ex ante payoff, τa_i. If the revolt succeeds, then those who participated receive the value of their full productive potential net of costs, $a_i - c_i$. Those who did not participate receive some lesser share of the their productive potential, ϕa_i, where $0 < \tau < \phi < 1$. The parameter ϕ captures the idea that veterans of a successful revolt are often privileged in the post-revolt era relative to those who never participated.

Given these parameters, the expected utility of participation is

$$\pi(P) = \omega(f_t)a_i - c_i,$$

while the expected utility of non-participation is

$$\pi(N) = [\omega(f_t)(\phi - \tau) + \tau]a_i.$$

Figure A.1 shows how payoffs for participation and non-participation vary over the probability of a successful revolt. The solid lines graph payoffs for an individual with high productive potential ($i = 1$), while the dashed lines graph payoffs for an individual with low productive potential ($i = 2$). We set the cost of participation to be equal for both individuals ($c_1 = c_2 = c$) so that we can see clearly the effect of changes in productive potential (a_i). For the individual with high productive potential, participation is preferred to non-participation so long as that individual believes that $\omega(f_t) > \omega_1^*$. For the low productive potential individual, no beliefs would lead to a preference for revolt, as $\pi_2(N) > \pi_1(P)$ for all values of $\omega(f_t)$, and so this individual is always "inframarginal" relative to the participation threshold. Wood (2003, pp. 270-274) discusses various

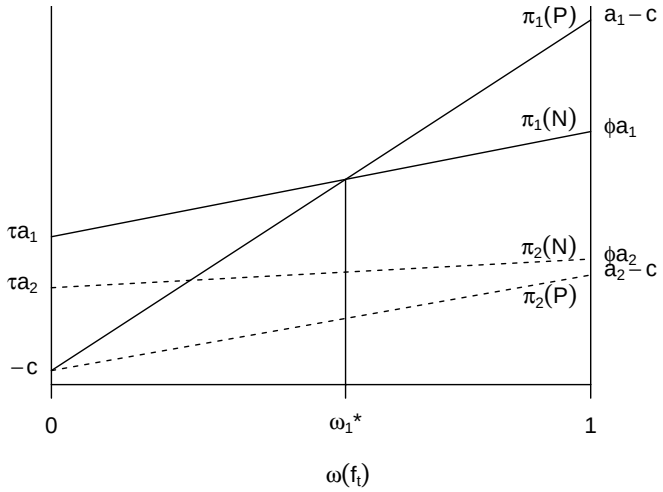


Figure A.1: Payoffs from participation (P) and non-participation (N) over the probability of a successful revolt ($\omega(f_t)$) for individuals with high productive potential ($i = 1$) and low productive potential ($i = 2$) but the same costs of participation (c).

processes through which beliefs about $\omega(f_t)$ might be shaped, such as insurgents publicizing their victories.

These primitives yield a coordination game with high participation and no participation equilibria (Medina, 2007; Wood, 2003). If an individual believes $\omega(f_t) > \omega_i^*$, then that individual will participate. Then in period $t+1$, we can assume that such participation will induce others to update their beliefs such that $\omega(f_{t+1}) > \omega(f_t)$. This makes it rational for those with $\omega(f_{t+1}) > \omega_j^*$ to want to participate, marking a step in a participation cascade (Granovetter, 1978; Kuran, 1989).

As Figure A.1 makes plain to see, the threshold value of $\omega(f_t)$ at which an individual determines that participation is preferable to non-participation is decreasing in productive potential. Call the threshold for an arbitrary individual ω_i^*. Then, for individuals with $0 < \omega_i^* < 1$,

$$\omega_i^* = \frac{\tau a_i + c_i}{a_i[1 - (\phi - \tau)]},$$

in which case,

$$\frac{\partial \omega^*}{\partial a_i} < 0.$$

This partial relationship will be determinative insofar as the costs of participation, c_i, do not covary strongly with a_i. As discussed above, c_i measures peaceful opportunities forgone when one decides to join. Many economic accounts of insurgency presume that material opportunity costs under the status quo will tend to override whatever gains the revolt might promise to bring about (e.g., Collier and Hoeffler, 2004; Miguel, Satyanath and Sergenti, 2004). However, in situations where repression is severe, status quo opportunities may vary only slightly from person to person, while productive potential will vary substantially. Then, those with high productive potential may have

only a very limited range of status quo opportunities that they would have to forgo by joining the revolt, and they may be attracted to the rewards that a successful revolt would offer (Medina, 2007). Our theoretical discussion leads to the following hypothesis:

> *H_a: Among harshly repressed groups (that is, groups for which τ is of substantial magnitude and c_i does not covary strongly and positively with a_i), participation will be increasing in productive potential (a_i).*

This stands the conventional economic analysis on its head, insofar as those with high productive potential are typically those whose status quo economic conditions appear favorable relative to their in-group peers. The difference is due to the interaction between the contextual and individual conditions.

This analysis focuses on a form of relative deprivation defined in terms of comparisons to a counter-factual of realizing one's own productive potential under less repression, rather than comparisons to other individuals (Olson and Roese, 2002). Knowing someone's position in the income distribution, say, would be inadequate to determine whether someone senses such deprivation. In order to determine levels of deprivation, it is necessary to measure the repressive constraints an individual faces and whether those constraints are visible enough to the individual—and attributable to the actions of certain groups—so as to provoke a comparison of what the individual feels she is capable of achieving versus what she is realistically *able* to achieve. Unfavorable counterfactual comparisons fuel resentment toward the custodians of the status quo. Individuals pressing hardest against the "glass ceiling" will have the most profound desire for political change, and thus have the strongest preferences for revolt. More immediately, individuals with high self-perceived productive potential are likely to see themselves as suited for leadership roles in the rebellion and under any new regime that it ushers in (VanBelle, 1996).[1]

The main macro-level scope condition for our hypothesis is the level of repression, which affects τ and the correlation between a_i and c_i. Under less oppressive circumstances, it may be argued that a class of elites with motivation to stir-up political change would hire lower-class members of their group as the labor behind a movement to overthrow the current regime (Esteban and Ray, 2011; Esteban, Mayoral and Ray, 2012). It is when repression is severe enough, meaning the individuals with high productive potential are not able to access sufficient resources under the status quo, that individuals with high preference for revolt would have ample motivation to join the revolt themselves, and may not even be able to finance the participation of others in the revolt.

[1] Chinodya (1990) provides a moving account of how opportunities afforded to those who participated in the Zimbabwean rebel movements translated into privileges and positions of power after independence and the assumption of majority rule.

B Burundi case background

Like neighboring Rwanda, the social-political system of Burundi is marked by the caste-like Hutu-Tutsi ethnic structure—an example of what Horowitz (1985) calls a "ranked" ethnic system. The political system in place by the early 1990s was an intensified version of a hierarchical ethnic system that had prevailed at least since colonial times (1903-1962) and most likely for a considerable period prior.[2] Over the decade following independence, power and access to elite opportunities were concentrated in the hands of a southern Tutsi elite. Hutu "liberationist" organizations emerged in the 1970s, including the *Parti pour la liberation du peuple hutu* (PALIPEHUTU) and its armed wing, the *Forces nationales de liberation* (FNL), as well as the *Front pour la liberation nationale* (FROLINA) and its armed wing, the *Forces armees du peuple*. While Tutsi elites dominated the national political system to the exclusion of Hutus, at the more local level and in non-elite spheres, some opportunity for Hutu mobility remained.

A Hutu uprising that took place in April 1972 led to a discontinuous increase in repression of Hutu opportunities and upward mobility. We use this tragic event as the basis of a cohort-based test of the effects of increased repression. We now provide relevant historical details to explain the cohort-based test.

Political unrest began in the southwest of the country and triggered the mutiny of certain Hutu troops in the army. The uprising has been variously attributed to Hutu militants originally associated with the Burundi Workers' Party, as well as renegade members of the Burundian gendarmerie, among other accounts (Lemarchand, 1994; Manirakiza, 1992; Reyntjens, 1994). Whatever the cause, the consequence was a genocidal crackdown implemented by the southern-Tutsi-dominated military. Death toll figures typically mentioned are between 100,000-200,000, almost exclusively Hutu (Ngaruko and Nkurunziza, 2000). The army's response in 1972 is sometimes characterized as a "decapitation" of Burundi's Hutu subpopulation.[3] The army's strategy was apparently based on a fear that upwardly mobile and ambitious Hutus would inevitably seek to overthrow the southern-Tutsi-dominated regime. Therefore, Hutu men in relatively high positions—teachers, doctors, bankers—were targeted for imprisonment or execution. This drove many Hutu intellectuals into exile, largely to Rwanda. It instilled fear among Hutu families that caused many to avoid sending their children to school, widening the gap in formal educational attainment across ethnic lines (Jackson, 2000). The generation of Hutus that grew up in the aftermath of 1972 faced intensified constraints on their ability to realize their ambitions through education or occupational mobility (Lemarchand, 1994, 106-117). In our empirical investigation, we analyze cohorts of men whose opportunities would have been most affected by this shock to see whether they exhibited a heightened proclivity to join the rebellion years later in 1993.

Moving ahead a few decades, the southern-Tutsi dominated government reacted to both internal and external pressures and initiated tentative steps toward democracy in the late 1980s. Pressures

[2] Lermarchand (1996) and Chretien (2003) provide descriptions of the historcal origins of ethnic and political divisions in the Great Lakes region, a topic that is subject to many disagreements.

[3] This characterization is based on authors' own interviews with journalists in Burundi, family-members of victims of the 1972 violence, and retired army officers who were active at the time of the massacre.

Table B.1: Estimated Sizes of Civil War Forces as of January 2004

Name of Force	Estimated Size
National army & armed police (Forces Armées Burundaises)	45,000
Rebel factions	
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

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

**Not a party to the peace process until September 2008 and so members were not included in the survey.*

and fears embedded in the ethnically divided system mixed toxically and made for a tumultuous democratization process. Competitive and relatively free elections were held in June 1993, which brought Melchior Ndadaye and his *Front pour la democratie au Burundi* (FRODEBU) to power by huge margins. Ndadaye became the first leader in Burundi's history that was simultaneously civilian, non-Southern, and Hutu. However, Ndadaye's term was abruptly cut off in October 1993 when he was assassinated in a bungled coup attempt by military officers from a hardline Tutsi faction. The coup attempt resulted in what an independent United Nations commission described as genocidal reprisals by Hutu mobs against Tutsi men in the countryside, followed by massacres of Hutus by the national army (United Nations, 1996).

As Alexis de Tocqueville wrote, "evils which are patiently endured when they seem inevitable become intolerable once the idea of escape from them is suggested" (de Tocqueville, 1856). This sentiment captures perfectly the moment for those who had voted Ndadaye into office, only to witness his assassination. Prominent Hutu politicians defected from the government to organize militant resistance to the re-consolidation of a Tutsi-led military oligarchy. These defectors organized politically under the banner of the *Conseil national pour la defence de la democratie* (CNDD), and soon established a military wing, the *Forces pour la defense de la democratie* (FDD). The CNDD and FDD claimed that their main goal was not "Hutu liberation" per se, but rather a return to "genuine democracy."

The FDD began to recruit people into the struggle to reinstate a political system that would allow the country's Hutu majority to exercise what its members saw as its fair share of power. The first guerrilla attacks attributed to the FDD took place in January 1994 in the suburbs of Bujumbura, spreading to other areas of the country. The media also recorded an escalation of attacks by the FNL over the ensuing year. Table B.1 shows the breakdown of participant numbers by factions throughout the 1993-2003 war.

Micro-level research by Humphreys and Weinstein (2008) on the Revolutionary United Front in Sierra Leone and Blattman (2009) on the Lord's Resistance Army in Northern Uganda has focused researchers' attention on the role of abduction in insurgent recruitment. Data from the

survey of Burundian rebels (described below) describe a very different situation from these two insurgent groups. While it is true that 67% of the ex-rebels that were interviewed indicated that they feared harm if they did not participate, 83% of these feared harm by the government, and only 28% of these feared harm by the rebels. (The numbers sum to more than 100% because the survey allowed people to indicate fear from both.) Only 7% indicated that they themselves had been physically mistreated by the rebel group prior to their becoming a member, and 15% indicated that someone in their family had been. This is in contrast to reports of mistreatment by government forces, which tally at 56% for respondents themselves and 73% for their families. More than half (51%) indicated that they had desired to join the rebel group for a while before they actually did so. When asked how they were convinced to join, 56% indicated that they convinced themselves, while 25% indicated that their friends convinced them. Finally, 66% indicated that they had friends in the group before they joined. The picture that these self-reports paint is not one where abduction plays a major role, but rather a combination of individual volition, peer pressure, and threats from government forces, similar to what Kalyvas and Kocher (2007) describe for Vietnam and Greece.

This brief historical summary serves to demonstrate that the Burundian civil war is a rich case for the study of the relationship between the structural condition of horizontal inequalities and individual-level motivations for violence. Hutus suffered severe constraints to their educational and professional opportunities under Tutsi minority rule.[4] Our theoretical analysis predicts that those Hutu men who had higher perception of their productive potential would have stronger motivation to revolt. Our basic empirical analysis below tests this prediction using measures of educational attainment as a proxy for self-perceived productive potential. The 1972 uprising led to a discontinuous ratcheting up of constraints on Hutu upward mobility. This allows us to construct a more refined test. The discontinuous increase in repression generated a cohort who had completed school, and thus had their productive potential most clearly revealed to them in the classroom, only to have doors to further opportunities slammed in their faces. Our theoretical analysis suggests that these individuals would have especially strong motivation to rebel.

[4] Jackson (2000) is a thorough account of obstacles that Hutus faced in the education system. Ngaruko and Nkurunziza (2000, pp. 389-391) discuss barriers to Hutus' access to employment in the public bureaucracy as well as other barriers to Hutu mobility. Lemarchand (1994, 106-117) describes the construction of a "Tutsi ethnocracy" in the aftermath of the 1972 violence.

C Survey design

As indicated in the main text, the survey sample was a stratified multi-stage sample that was meant to be representative of both the ex-combatant and non-combatant populations as of 2007. Burundi is divided into seventeen provinces, each of these provinces is divided into communes and each of these communes is divided into collines, the smallest administrative unit in Burundi. The first stage of civilian sampling was conducted by randomly selecting communes that were stratified according to size. Then, within each commune, three-quarters of the civilian surveys were conducted at the commune center, and the remaining surveys were conducted in randomly selected outlying collines where the population was more sparse. When the collines were sufficiently small, households were numbered and then randomly selected. In the larger collines and in the city centers, households were randomly selected by having enumerators walk for a specified interval of time, and then surveying the nearest household. Finally, each household respondent was randomly chosen from among a list of all adults residing in the household.

The sample of ex-rebel combatants includes both those who had been demobilized as well as those who were integrated into the new army and police. A minority of ex-rebels were integrated into the new army and police, and the rest were put into the national disarmament, demobilization and reintegration (DDR) program. This selection was dominated by the CNDD-FDD and also involved technical vetting by existing army members and the United Nations. That is why it is crucial that our sample contains both types of ex-rebel combatants. The sample of demobilized combatants was randomly selected from the national DDR program's administrative list, stratified by former group affiliation and rank. Note that inclusion on this list was subject to negligible amounts of self-selection. This is because the material benefits being offered were quite attractive, and there were only negligible numbers of demobilized rebels who could afford to forgo such benefits (Gilligan, Mvukiyehe and Samii, 2013). The sample of army and police members was drawn at random from barrack personnel lists, with the selected barracks being any and all barracks present in the communes selected for the civilian sample. The analyses below use sampling weights to account for the fact that sampling rates differed from commune-to-commune (depending on the population sizes as well as sampling targets) and also to account for the "case-control" or "response-based" sampling design, whereby we purposefully sampled on the basis of rebel participation status, the outcome of interest (Korn and Graubard, 1999, Ch. 9; Manski, 1995, Ch. 4).

The survey sample was a stratified multi-stage sample that was meant to be representative of both the ex-combatant and non-combatant populations as of 2007. The survey was multipurpose, and therefore asked questions about family background; living conditions prior, during, and after the war; participation in armed groups during the war; motivations for joining and perceived aims of the armed groups; experiences with armed groups and third party intervening actors; and attitudes toward post-conflict policies and conditions.[5] Burundi is divided into seventeen provinces, each of these provinces is divided into communes and each of these communes is divided into collines, the smallest administrative unit in Burundi. The first stage of civilian sampling was conducted by randomly selecting communes that were stratified according to size. Then, within each commune, three-quarters of the civilian surveys were conducted at the commune center, and the

[5] The full survey instrument is available at [[SITE ADDRESS TO BE ADDED]].

remaining surveys were conducted in randomly selected outlying collines where the population was more sparse. When the collines were sufficiently small, households were numbered and then randomly selected. In the larger collines and in the city centers, households were randomly selected by having enumerators walk for a specified interval of time, and then surveying the nearest household. Finally, each household respondent was randomly chosen from among a list of all adults residing in the household.

The sample of ex-rebel combatants includes both those who had been demobilized as well as those who were integrated into the new army and police. A minority of ex-rebels were integrated into the new army and police, and the rest were put into the national disarmament, demobilization and reintegration (DDR) program. This selection was dominated by the CNDD-FDD and also involved technical vetting by existing army members and the United Nations. That is why it is crucial that our sample contains both types of ex-rebel combatants. The sample of demobilized combatants was randomly selected from the national DDR program's administrative list, stratified by former group affiliation and rank. Note that inclusion on this list was subject to negligible amounts of self-selection. This is because the material benefits being offered were quite attractive, and there were only negligible numbers of demobilized rebels who could afford to forgo such benefits (Gilligan, Mvukiyehe and Samii, 2013). The sample of army and police members was drawn at random from barrack personnel lists, with the selected barracks being any and all barracks present in the communes selected for the civilian sample. The analyses below use sampling weights to account for the fact that sampling rates differed from commune-to-commune (depending on the population sizes as well as sampling targets) and also to account for the "case-control" or "response-based" sampling design, whereby we purposefully sampled on the basis of rebel participation status, the outcome of interest (Korn and Graubard, 1999, Ch. 9; Manski, 1995, Ch. 4).

Because the data are from a retrospective case-control survey, there are certain vulnerabilities to bias that must be addressed. First, there is the possibility of selection bias due to non-random survivorship during the war. We address this issue by controlling for survivorship probability. Another section of this Appendix contains details on differential survival and our methods for addressing it. Briefly, our survey data include rosters of those killed during the war. We use the data from these rosters to estimate survival probability for Hutu men age 13-42 in 1993 and then control for this survival probability in our analysis.[6] A second issue is that retrospective surveys can suffer from the fact that individuals may have poor recall, which can result in noisy measurements. However, given that this study focuses on objective information, such as the respondent's father's level of education, the risk for this type of bias is mitigated.[7]

[6] We use the linear predictor (or log-odds) from a penalized logistic regression of survivorship on home province fixed effects, father's education, family size, age, and rebel participation status, accounting for survey design features, using a flexible specification.

[7] Questions of this objective nature can be contrasted with survey questions that ask a respondent to remember his or her attitudes or feelings at a certain point in time, which potentially suffer from a much higher concern for recall bias.

D Survival and attrition

In this section, we describe issues related to survival and the controls that we introduce to account for differential survival probabilities that we observe. We focus on the population of interest for this study—namely, the living population of Hutu males who were between 13 and 42 years of age at the onset of the civil war in 1993. The survey administered to both civilians and those who joined armed groups included a roster that recorded information on all of the the respondents' family members who died between 1993 and the administration of the survey in 2007. The roster also collected information on whether or not the deceased individuals had joined an armed group prior to their death. The roster allowed us to gather data on 488 Hutu males that were aged 13 to 42 years in 1993 but who passed away prior to 2007, of whom 76 had joined the rebellion prior to their death. Note that these figures do not account for the survey weighting, e.g., to account for higher rates of sampling among armed group participants relative to civilians. They merely represent the raw tabulation of the survey data.

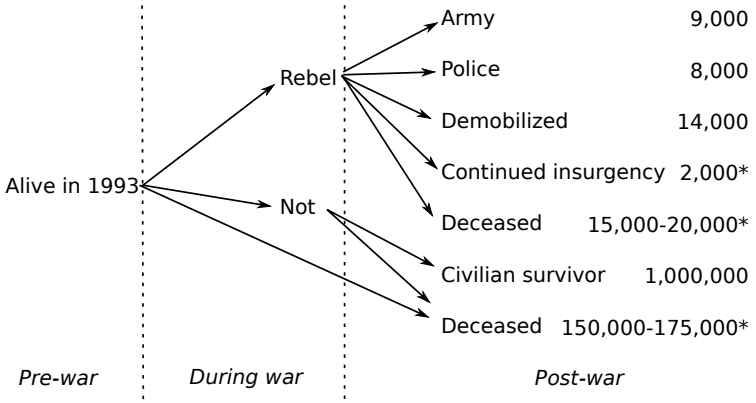


Figure D.1: Wartime and post-war life trajectories for Hutu men aged 13 to 42 in 1993. Figures to the right are estimates on the numbers of men who were in each of the seven listed outcome states as of 2007. Figures with asterisks estimates based on rosters collected in the survey that gathered information on those who died during the war. Sources: authors' calculations from survey data and Gilligan, Mvukiyehe and Samii (2013, Suppl. Appendix).

Using our various data sources, we can construct a life trajectory diagram for Hutu males above the aged 13 to 42 in 1993. This combines information from our roster on deaths, our survey data on armed group participation, census figures for 1993, the data on the numbers of individuals who joined the various armed groups from Table B.1 in the main text, and finally conventional estimates for the size of the FNL-PALIPEHUTU forces as of 2007. This diagram is displayed in Figure D.1. Starred values in the figure are for those types of individuals that the survey could not sample directly. Our rosters provide proxy information for the deceased. The rosters suggest that the number of deceased was very large. This generally corresponds to conventional estimates used by international organizations such as the United Nations and World Bank, who put the combined

death toll for the entire population at about 300,000.[8]

These deaths represent a form of attrition for our analysis of the 1993 male population. Such attrition could bias our analysis. Specifically, suppose that the probability of death were relatively high among either (i) rebel participants with low levels of education for themselves or their fathers, or (ii) non-participant civilians with high levels of education for themselves or their fathers. Either of these patterns would confound the 2007 relationship between rebel participation and education. Even if there were no causal relationship between education and rebel participation, either of these patterns could induce a spurious positive correlation between rebel participation and education.

The data indicate that such patterns are in fact present. This means that it is important to control for survival probability in our analyses (and in fact we do for the tables presented in the main text). Because the deaths data were collected for siblings of respondents, we know the level of fathers' education for each of the deceased. We do not have data on the level of educational attainment of the deceased. In any case, this would be a censored variable for those who died prior to completing their education. As such, we can conduct an empirical analysis of death rates by father's education, distinguishing between those who joined the rebellion and those who did not. The results of this analysis are shown in Table D.1. The first column shows results of a regression of survival on father's education, whether or not the individual joined the rebellion (recall that we have this information even for the deceased), and then their interaction. (The analysis include home province fixed effects, along the lines of all analyses presented in the main text.) We see that the probability of death is much higher for rebels than civilians, but that this probability declines (by about 6.5 percentage points) for each year of father's education. For civilians, the correlation between father's education and the probability of death is negligible. The second column introduces controls, including age and the number of brothers that someone had. The reason that we control for the number of brothers is because the likelihood that a death would be picked up by the survey is a function of the number of brothers available to be interviewed. We see that the patterns from the first column do not change appreciably upon introducing controls.

These findings indicate that we do indeed need to account for probability of survival so as to reduce the potential for estimating a spurious relationship between education and rebel participation. To do so, we apply the methods proposed by Angrist (1997) and Newey (2009), which is to estimate the probability of survival, and then to control for this probability in one's analysis. We estimate survival probability using age, father's education, number of brothers, whether the individual joined the rebellion or not, and home-province dummy variables. We also include the survey sampling probabilities into the model, given that they contain information relevant to the sampling strata (see the discussion on sampling in the main text). We use a flexible specification, including quadratic terms for all continuous variables and then all two-way interactions. Because we are interested in the predicted values, rather than the coefficients, we use a penalized logistic regression, which is known to have better predictive performance in situations where one includes many higher order terms, interactions, and dummy variables.[9] The coefficients from the penalized

[8] See, e.g., United Nations News Service, 2007, "Burundi faces hurdles despite progress, Security Council told," *UN Daily News*, p. 6, available at http://www.un.org/news/dh/pdf/english/2007/21052007.pdf

[9] We fit the model using the `bayesglm` function in the `arm` package for R. See Gelman et al.

Table D.1: Death as a function of father's education, rebel participation, and control variables

	Model 1	Model 2
(Constant)	0.333^{***}	0.243^{*}
	(0.089)	(0.104)
Father's ed. (yrs.)	0.009	0.010
	(0.006)	(0.006)
Rebel participant	0.621^{***}	0.550^{**}
	(0.093)	(0.169)
Father's ed. X rebel	-0.065^{***}	-0.064^{***}
	(0.020)	(0.019)
Age		0.003
		(0.003)
No. brothers		0.024^{*}
		(0.011)
Age X rebel		-0.005
		(0.010)
No. brothers X rebel		0.015
		(0.033)
N	1393	1393
R^2	0.134	0.147
adj. R^2	0.120	0.132
Resid. sd	9.662	9.599

Weighted least squares with standard errors accounting for clustering at the commune level.
All models account for pre-war province fixed effects.
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

logistic regression fit are displayed in Table D.2.

As a technical note, we find that while number of brothers is strongly predictive of survival probability, it does appear to have a strong relationship to rebel participation after accounting for the other factors that we consider in our main specifications in the main text. These results are presented in the first and third results columns in Table E.4 of the appendix to the main text. (Those results are in a separate section because they are part of a separate discussion that concerns number of brothers as a predictor of participation.) This provides a source of plausibly exogenous variation in our survival probability model, which means that our ability to control for survival probability is not dependent solely on non-linearities in our specification for the model presented in Table D.2 (Angrist, 1997).

In our regressions, we control for survival probability by using the linear predictor (or log-odds) from the penalized logistic regression. This provides a robust semi-parametric approach to controlling for selection bias (Angrist, 1997; Newey, 2009).

(2008).

Table D.2: Modeling survival probability (penalized logistic regression coefficients)

	Model 1
(Constant)	-0.242
	(0.347)
Age	-0.030
	(0.024)
Age sq.	0.003^{*}
	(0.001)
Father's ed.	$-0.107^{\dagger}$
	(0.064)
Father's ed. sq.	0.010^{*}
	(0.005)
Sampling weight	0.002^{***}
	(0.000)
Sampling weight sq.	-0.000^{***}
	(0.000)
No. brothers	-0.303^{**}
	(0.114)
No. brothers sq.	0.016
	(0.011)
Age X father's ed.	-0.004
	(0.003)
Age X sampling weight	-0.000
	(0.000)
Age X No. brothers	-0.004
	(0.005)
Father's ed. X sampling weight	-0.000
	(0.000)
Father's ed. X no. brothers	-0.007
	(0.011)
Sampling weight X no. brothers	0.000
	(0.000)
Rebel participant	3.064^{***}
	(0.395)
Rebel X Age	0.016
	(0.023)
Rebel X father's ed.	0.114^{*}
	(0.051)
Rebel X no. brothers	-0.041
	(0.081)
Rebel X sampling weight	-0.008^{***}
	(0.002)
N	1393

Standard errors in parentheses
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

E Checking alternative explanations

We explore five potential alternative explanations for the positive correlation among Hutu males between participation in the insurgency and one's father's and own education. The first alternative explanation has to do with possible revenge motivations based on the fact that the 1972 crackdown targeted Hutu elites. The second is based on the possibility that rebel movements selectively recruited more educated people, as proposed by Bueno De Mesquita (2005) in a review of evidence from Krueger (2007)'s similar findings regarding the positive correlation between socio-economic status and participation in Islamic militancy. The third is based on the possibility that *household* opportunity costs may have been lower for higher-education households, and that our results are spurious to such household opportunity costs. The fourth is based on the idea that education yields resources—such as skills or networks—that facilitate collective action in ways that are distinct from economic opportunities. The fifth is that education is a gateway to political socialization that can motivate participation in ways that are distinct about grievances regarding economic opportunity. This is how Krueger (2007) interprets his own findings.[10] Figure E.1 is a causal diagram that maps these various alternative explanations along with the "perceived productive potential" mechanism that is emphasized by the glass ceiling theory. We see that one's father's and own education can each affect self-perceived productive potential, possibly in ways that reinforce such self-perceptions over time, but that education can trigger other mechanisms that could affect participation.

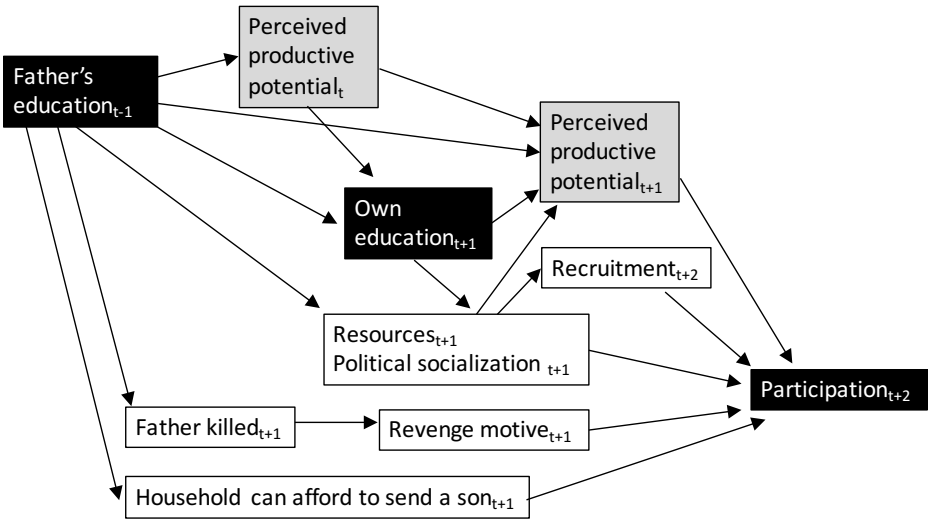
The sections that follow consider these five alternative explanations. Evidence in favor of any of these alternatives would not necessarily render the glass ceiling logic irrelevant, but it would raise questions about its preeminence for interpreting the positive correlation between education and participation. That said, as detailed below, we do not find compelling evidence for these alternatives over and above the glass ceiling logic. This does not prove the truth of the glass ceiling logic, but it does mean that the glass ceiling logic remains as a parsimonious explanation for the various empirical patterns analyzed here (Clarke and Primo, 2012). Table E.7 below gives a summary overview of the findings.

E.1 Revenge motivations

Higher rates of participation among more educated individuals could have been based on revenge motivations, given that the Tutsi-dominated military regime targeted educated Hutus in the crackdown that followed the 1972 Hutu uprising. If the fathers of more educated individuals were more likely to be killed, then these more educated individuals may have been motivated by the desire to avenge their fathers' deaths rather than frustrations precipitated by pushing against the"glass ceiling." We do not have detailed information on who in our sample, exactly, was targeted in 1972. But we do have information on whether respondents' fathers were alive in 1993. Thus, as an indirect test we fit both the basic equation (1) and cohort-specific equation (3) using as the outcome the indicator for whether the respondent's father was alive in 1993. Confounding would appear as significant negative coefficients on β_1 and α_3.

[10] We thank anonymous reviewers for proposing the other potential alternative explanations.

Figure E.1: Causal diagram illustrating potential alternative explanations



The figure shows various causal pathways through which father's and own education may cause participation in the insurgency. The black boxes show the variables that we have measured and include in the statistical analysis. The gray boxes show the causal mechanism that involves "perceived productive potential," which is the mediating variable that is highlighted by the glass ceiling theory. The white boxes show alternative channels that we attempt to address in the analyses in this section.

The survey also asked rebel recruits their reasons for joining the insurgency. As a second attempt to test this alternative revenge hypothesis, we use the sample of rebel recruits to fit both the basic equation (1) and cohort-specific equation (3), where the outcome is whether the respondent gave "revenge" as one of their stated reasons for joining the rebellion. If this alternative hypothesis were valid, father's education would be positively correlated with this revenge response, especially among older cohorts.

Tables E.1 and E.2 show results for these tests of vengeance motivations. We find no evidence that father's survival to 1993 was decreasing in father's education levels (Table E.1, Model 1). Neither are there negative cohort effects (Models 2 and 3). Finally, recall that in Table 2 in the main text we found that one's own educational attainment was positively associated with participation even after controlling for father's education. To the extent that vengeance motives were the primary driver, we would expect father's education to be the primary driver of the effect, and the apparent effect of one's own education would be spurious. This does not appear to be the case. If anything, fathers' survival probabilities are increasing in education, and this positive relationship is stronger for older cohorts. As Table E.2 in the appendix shows, none of the productive potential variables significantly predict a response that a rebel recruit participated for revenge purposes. At first glance this may appear as puzzling given that the historical record suggests the targeting of high status individuals in 1972. However, the 1972 killings by the regime were largely concentrated in the

Table E.1: Father alive in 1993 as a function of father's education, by age categories in 1972

	Model 1	Model 2	Model 3
(Constant)	0.920^{***}	0.908^{***}	0.918^{***}
	(0.097)	(0.091)	(0.100)
Age in 1972	-0.026^{***}	-0.026^{***}	-0.026^{***}
	(0.005)	(0.005)	(0.005)
Age in 1972 sq.	0.000	0.001^{**}	0.001
	(0.000)	(0.000)	(0.001)
Father's ed. (yrs.)	0.003	0.000	0.002
	(0.008)	(0.008)	(0.008)
Age in 1972 > 12		-0.306^{*}	
		(0.153)	
Father's ed. (yrs.) X Age in 1972 > 12		0.038	
		(0.034)	
Age in 1972 > 16			-0.110
			(0.161)
Father's ed. (yrs.) X Age in 1972 > 16			0.053
			(0.034)
N	905	905	905
R^2	0.250	0.260	0.251

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

(Number of clusters is 65.)

Dependent variable is father alive in 1993 (0,1).

All models control for pre-war province fixed effects and survival probability.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

‡ jointly significant at $p < .10$.

southwestern and central portions of the country, in which case the killings targeted a specific segment of Hutu men, while the increased repression applied to Hutus throughout Burundi. For this reason, when looking at national-level patterns, it is quite possible that father's survival was generally increasing in parts of the country outside the areas targeted by the regime in 1972, and that the sons of high attainment fathers suffered the consequences of a nationwide repression of Hutu mobility.

E.2 Selective recruitment

For the second alternative explanation, we use data from our sample of Hutu men who *did not* participate in the insurgency. The survey asked these men whether they had *considered* joining. To test the selective recruitment explanation, we run our basic equation (1) on this civilian sample, using their responses to this question as the outcome. If selective recruitment was occurring, we

Table E.2: Participation for "revenge" as a function of father's education and by age categories in 1972

	Model 1	Model 2	Model 3
(Constant)	0.122^{*}	0.116^{*}	0.126^{*}
	(0.052)	(0.053)	(0.053)
Age in 1972	0.004	0.005	0.004
	(0.003)	(0.003)	(0.003)
Age in 1972 sq.	$-0.000^{\dagger}$	-0.000	-0.001
	(0.000)	(0.000)	(0.000)
Father's ed. (yrs.)	0.006	0.007	0.007
	(0.006)	(0.006)	(0.006)
Age in 1972 > 12		-0.102	
		(0.136)	
Father's ed. (yrs.) X Age in 1972 > 12		-0.016	
		(0.010)	
Age in 1972 > 16			0.130
			(0.163)
Father's ed. (yrs.) X Age in 1972 > 16			-0.021
			(0.015)
N	446	446	446
R^2	0.050	0.053	0.053

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

(Number of clusters is 65.)

Dependent variable is participation in armed revolt (0,1).

All models control for pre-war province fixed effects and survival probability.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

would expect the coefficients on education and father's education to be negative, since this would indicate that undereducated Hutu civilians considered joining but never ended up participating, possibly due to selective recruitment. As the results in Table E.3 in the appendix show, one's own education is positively correlated with having considered joining the rebellion (significant at the 10% level), which indicates that in fact there were a significant number of *well-educated* civilians who considered joining the movement but never ended up participating. This contradicts the alternative explanation that rebel leaders selectively recruited based on educational attainment. If recruitment had been selective on education, we would find a concentration of relatively low-educated men in our civilian sample who wanted to join, implying a negative coefficient. Father's education is not significantly correlated with having considered joining, although the sign on the coefficient is negative. Aside from the statistical imprecision of the estimate, this evidence does not strike us as compelling since it is difficult to image that recruitment would targeted subjects' fathers' education more than subjects' own education.

Table E.3: Whether civilian considered joining, as a function of own and father's education

	Model 1	Model 2
(Constant)	0.068	0.142
	(0.148)	(0.151)
Age in 1972	-0.019^{**}	-0.020^{**}
	(0.006)	(0.007)
Age in 1972 sq.	0.001	0.001
	(0.001)	(0.001)
Subject's ed. (yrs.)	$0.012^{\dagger}$	
	(0.007)	
Father's ed. (yrs.)		-0.006
		(0.011)
N	459	459
R^2	0.193	0.187

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

(Number of clusters is 61.)

Dependent variable is whether civilian considered joining (0=no,1=yes).

All models control for pre-war province fixed effects and survival probability.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

E.3 Household opportunity costs

The third alternative explanation proposes that we consider household-level opportunity costs, rather than opportunity costs that pertain only to individual would-be recruits. The argument here is essentially one of omitted variable bias: the relationship between education and participation in rebellion may be spurious as an indication of the effects of "productive potential." Rather, the argument goes, households with more educated fathers and more educated sons may be wealthier and therefore more able to "afford" a son to join the rebellion. First, it should be noted that this is quite a departure from the conventional "opportunity costs" account, which focuses on the material opportunity costs to the individual recruit. For this alternative account to have traction, it must be that would-be recruits are either altruistic toward their household members, or heads of households can dictate what their sons do.

We examined two measures of household opportunity costs that fit this alternative logic: (1) an index of prewar household wealth and (2) the number of brothers that the would-be recruit had. Each of these should affect the extent to which a household could "afford" one of its sons joining the rebellion. The index of pre-war wealth is the sum of "yes" responses to questions asking about whether the respondent's household owned land, cows, a radio, and beds with bed sheets.[11] If the findings with respect to education are indeed spurious, then controlling for these

[11] The last two items were selected based on focus group discussions, which indicated that these factors would be highly discriminating of households in terms of their wealth in 1993, while

Table E.4: Evaluating whether factors related to household opportunity costs confound estimates of the relationship between one's own education and participation

	DV=No. bros.	DV=Rebel	DV=Wealth	DV=Rebel	DV=Rebel
(Constant)	3.694^{***}	0.042^{*}	1.929^{***}	0.030	0.029
	(0.696)	(0.021)	(0.157)	(0.021)	(0.021)
Age in 1972	-0.009	-0.001^{*}	0.002	-0.001^{*}	-0.001^{*}
	(0.017)	(0.000)	(0.014)	(0.000)	(0.000)
Age in 1972 sq.	-0.001	0.000	0.002	0.000	0.000
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
Subj. ed.	0.050	0.003^{***}	0.078^{***}	0.002^{**}	0.002^{**}
	(0.042)	(0.001)	(0.023)	(0.001)	(0.001)
Father's ed.	0.004	$0.002^{\dagger}$	0.059^{*}	0.002	0.002
	(0.040)	(0.001)	(0.028)	(0.001)	(0.001)
Number of brothers in 1993		0.000			0.000
		(0.001)			(0.001)
Wealth index				0.007^{*}	0.007^{*}
				(0.003)	(0.003)
N	905	905	905	905	905
R^2	0.112	0.026	0.178	0.028	0.028
adj. R^2	0.089	-0.000	0.156	0.002	0.001
Resid. sd	48.175	4.252	27.804	4.248	4.250

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

(Number of clusters is 65.)

All models control for pre-war province fixed effects and survival probability.

In results columns 1 and 3, subj. ed. and father's ed. are jointly significant with p < .05.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

factors should nullify the relationship between educational attainment and rebellion. Table E.4 shows the results of this test. Results columns 1 and 3 show that the education measures are positively correlated with both the number of brothers and wealth (in both cases, subjects' own and father's education are jointly significant). We also find that participation probability is increasing in household wealth, although not so for number of brothers.[12] However, what is most important is that the relationship between participation and either one's own educational attainment or one's father remains very similar to what we showed in the main text, even after controlling for number of brothers and pre-war wealth. The coefficient on subject's own education is attenuated only somewhat (from 0.003 to 0.002 change in participation probability per year of education).

being very easy for respondents to recall.

[12] This positive coefficient on wealth provides further empirical indication that wealthier households were not paying poorer individuals to join the rebellion on their behalf.

E.4 Collective action resources

Insurgent organization requires collective action, and collective action is facilitated by various individual-level resources–for example, network ties and access to organizing information. The function of these resources are captured in our formal model above, in particular in the ways that they affect beliefs about others' participation (i.e., beliefs about f_t). Perhaps the primary role of father's and own education is to increase networks and access to information, which in turn may make it easier to learn about others' willingness to participate (f_t in the model), even if there is no effect on self-perceived productive potential (a_i in the model)? We assess this possibility by examining a few additional variables from the survey. The first is whether the family possessed a radio prior to war, which would facilitate access to collective-action relevant political information. The second is whether respondents reported that someone from a rebel faction had ever told them about the possibility of joining the group, which would be indicative of access to the insurgent recruitment network. The third is whether the respondent tended to spend most time with friends or with family prior to the war, which is intended to get at whether the respondent was active in peer networks that were more likely to be the social settings for political engagement and insurgent participation.

Table E.5 displays the results of statistical analyses as to whether these factors cofound our interpretation of the relationship between education and participation. For these information- and network-related factors to be confounders we should see that (i) they are predicted by the education variables and that (ii) controlling for these factors in the analysis weaken the correlation between the education variables and insurgent participation. Results columns 1, 3, and 5 show that the education variables do indeed positively predict radio access, whether one was informed about joining a rebel group, or tended to spend time more time with peers. However, results columns 2, 4, and 6 show that including these factors into the model that predicts insurgent participation does not affect the coefficients on the education variables.

E.5 Political socialization

The final alternative that we consider is that one's father's and own education may increase access to opportunities for political socialization. Indeed, Krueger (2007) interprets his own finding that militants tended to have higher education and socio-economic status in this way. The argument is that educational institutions themselves are sites of political socialization, and that education itself allows one to read and interpret political dynamics more clearly. Of course, this leaves open the question as to why militancy or insurgency would be the way that educated people engage in politics. Nonetheless, given the weight put on this line of thinking in the existing literature, we examine the potential for such socialization to confound our interpretation of the positive association between education and insurgent participation. The survey asked respondents which political party they had supported in 1993. The question was asked in a way that asked about "support," not about voting, because many of the respondents in our sample would have been too young to vote in 1993. The available responses were to support the incumbent and predominately Tutsi UPRONA party, the challenger and Hutu-dominated FRODEBU party, none of these parties, or some "other" party. Only a small number (26 respondents in total) reported "other," and when asked for more

Table E.5: Evaluating whether collective action resources confound estimates of the relationship between education and participation

	DV=Radio	DV=Rebel	DV=Informed	DV=Rebel	DV=Friends	DV=Rebel
(Constant)	0.649^{***}	0.042^{*}	0.365^{***}	0.042^{*}	0.364^{***}	0.043^{*}
	(0.066)	(0.020)	(0.110)	(0.020)	(0.074)	(0.020)
Age in 1972	-0.002	-0.001^{*}	-0.003	-0.001^{*}	-0.016^{*}	-0.001^{*}
	(0.006)	(0.000)	(0.010)	(0.000)	(0.007)	(0.000)
Age in 1972 sq.	$0.001^{\dagger}$	0.000	-0.000	0.000	$0.001^{\dagger}$	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Subj. ed.	0.023^{**}	0.003^{**}	0.014	0.003^{**}	0.043^{***}	0.003^{**}
	(0.008)	(0.001)	(0.011)	(0.001)	(0.009)	(0.001)
Father's ed.	0.004	$0.002^{\dagger}$	-0.002	$0.002^{\dagger}$	0.015	$0.002^{\dagger}$
	(0.009)	(0.001)	(0.012)	(0.001)	(0.010)	(0.001)
HH had radio pre-war		0.003				
		(0.004)				
Informed by faction pre-war				0.005		
				(0.005)		
Spent time w. friends pre-war						0.002
						(0.004)
N	905	905	905	905	905	905
R^2	0.133	0.026	0.095	0.027	0.196	0.026
adj. R^2	0.111	-0.000	0.072	0.000	0.174	-0.000
Resid. sd	11.886	4.252	14.425	4.252	13.346	4.252

Weighted least squares estimates.
Standard errors account for clustering at the commune level.
(Number of clusters is 65.)
All models control for pre-war province fixed effects and survival probability.
In results columns 1, 3, and 5, subj. ed. and father's ed. are jointly significant with p < .05.
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

details, these respondents overwhelmingly reported supporting the more radical Hutu-liberationist PALIPEHUTU party. So, we can take "other" responses as proxies for more radical political positions. Bear in mind that the survey was carried out in 2007, so these are retrospective self-reports. Assuming there are no systematic biases in the ways that respondents answered this question, we want to check whether (i) more educated respondents were either *less* likely to indicate support for "none" (implying a political socialization effect) or *more* likely to support "other" (implying a radicalization effect) and (ii) whether controlling for these political support variables reduces the strength of the relationship between the education variables and insurgent participation.

Table E.6 shows the results of analyses to check for such patterns. Results columns 1, 2, 3, and 4 offer no indication of either a political socialization or radicalization effect. Result column 5 shows that controlling for the political support variables does not weaken the association between the education variables and participation.

Table E.6: Evaluating whether political socialization confounds estimates of the relationship between education and participation

	DV=Sup. UPRONA	DV=Sup. FRODEBU	DV=Sup. none	DV=Sup. other	DV=Rebel
(Constant)	0.158^{*}	0.640^{***}	0.180	0.002	0.029
	(0.072)	(0.097)	(0.130)	(0.008)	(0.020)
Age in 1972	0.016^{**}	0.030^{***}	-0.047^{***}	0.001	-0.001^{*}
	(0.005)	(0.006)	(0.007)	(0.001)	(0.001)
Age in 1972 sq.	-0.001^{*}	-0.002^{**}	0.002^{***}	-0.000	$0.000^{\dagger}$
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Subj. ed.	-0.004	$-0.019^{\dagger}$	$0.020^{\dagger}$	0.001	0.003^{***}
	(0.006)	(0.011)	(0.011)	(0.001)	(0.001)
Father's ed.	0.011	-0.013	0.004	-0.001	0.003^{*}
	(0.013)	(0.009)	(0.013)	(0.002)	(0.001)
1993 support for FRODEBU					0.020^{***}
					(0.004)
1993 support for none					0.008
					(0.006)
1993 support for other					$0.126^{\dagger}$
					(0.071)
N	905	905	905	905	905
R^2	0.135	0.174	0.315	0.045	0.034
adj. R^2	0.113	0.152	0.297	0.020	0.006
Resid. sd	10.309	13.585	12.542	2.723	4.240

Weighted least squares estimates.
Standard errors account for clustering at the commune level.
(Number of clusters is 65.)
All models control for pre-war province fixed effects and survival probability.
In results columns 1, 2, and 3, subj. ed. and father's ed. are jointly significant with p < .05.
† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

E.6 Additional considerations

More complex causal stories are possible, but we do not think that they necessarily run contrary to our interpretation. Lee (2011), for example, finds that participation in Bengali militancy is generally increasing in socio-economic status, but that such is also true of participation in non-violent political activism. When one compares militants to non-violent activists, one finds that militants tended to be of relatively lower socio-economic status. Such a pattern is perfectly consistent with our theoretical account when the option of non-violent activism is viable. But this may not always be the case, and to the extent that such political activism is in pursuit of changes of the status quo, a situation such as what we document in Burundi may be expected. Kalyvas and Kocher (2007) point to the fact that in many insurgencies, individuals are forced to join due to threats of victimization by the government. Thus, one might propose that it was the fact that more educated Hutus were targeted by the regime that drove them to participate, rather than the "glass ceiling" logic. Such a proposition merely raises the question, *why* would more educated Hutus be targeted in the first place? A likely answer is that the regime *anticipated* that educated Hutus would be the most likely to organize to fight against the repressive status quo. Most likely, officials proposing to target educated Hutus had in their minds something akin to the glass ceiling logic.

A lingering puzzle is why more educated Hutus did not simply seek to leave rather than fight against the repressive regime. Our sense is that the option to flee was available to only a very small number of highly privileged individuals. Given the extent of repression, such an "exit" option would not have provided enough of a safety valve for the regime. And even for those who left, it

is worth reiterating from our contextual discussion in this Appendix that a number of Hutu leaders who agitated for change had spent time abroad, only to return to Burundi to participate in the fight at home.

Our study focuses primarily on socio-economic factors that relate to insurgent participation. This is not to dismiss the role of religion, culture, or other symbolic factors that may motivate participation (Atran, 2010). We feel that a focus on socio-economic factors is important because of the direct connection between such motivations and the types of economic policy levers that are often applied to mitigate conflict (Berman et al., 2011). Our analysis also seeks to investigate micro-foundations for the important macro-level correlation between horizontal inequality and conflict. A thorough investigation of other, non-economic factors would require much more space than we have here. Nonetheless we encourage complementary research along those lines.

E.7 Summary of key results and assessments of alternatives

Table E.7 provides an overview of the main results from the paper as well as results from statistical checks of potential alternative explanations contained in this appendix. Assessing the totality of the results, the glass ceiling logic remains as a parsimonious and robust explanation for these patterns.

Table E.7: Summary of empirical tests and interpretations

Table Number	Evidence	Interpretation
Table 2	Own and father's education predicts insurgent participation.	Shows that those with higher productive potential are more likely to revolt.
Table 3	Relationship between father's education and own accomplishments much weaker for Hutu as compared to Tutsi males.	Validates the assumption that Hutu men were subject to repression that limited their ability to realize their productive potential, as measured by their father's accomplishments.
Table 4	Father's education predicts own accomplishments more strongly for Hutu men after the war as compared to before.	Validates use of father's education as a proxy for productive potential and shows that Hutu men were indeed subject to repression that was then lifted as a result of the war.
Table 5	For insurgents, own education and father's education predict outgoing rank more strongly than incoming rank.	Validates use of own and father's education as a proxies for productive potential, and shows that those with high productive potential enjoyed mobility within the insurgency, providing a selective incentive.
Table 6	Father's education predicts own education more strongly for cohorts above school age in 1972, as compared to younger cohorts.	Shows that opportunities to pursue education were indeed sharply constrained after 1972, meaning that older cohorts were more likely to cultivate and recognize their own productive potential.
Table 7	Father's education predicts own participation in revolt more strongly for cohorts above school age in 1972, as compared to younger cohorts.	Shows that older cohorts who had relatively greater opportunity to cultivate, only to have doors to mobility slammed in their faces, were especially motivated to revolt.
Tables E.1 and E.2 (appendix)	Positive and insignificant correlation between father's education and father's survival; positive and insignificant cohort effect. Small and insignificant correlation between father's education and indication of "revenge" as motivation for participation among insurgents; negative and insignificant cohort effect for indicating "revenge" as a motivation for participation.	Inconsistent with alternative explanation of participation in revolt being motivated by grief or a desire for revenge over fathers being killed.
Table E.3 (appendix)	Positive and marginally significant correlation between own education and civilians' consideration of joining, and negative but insignificant correlation between father's education and civilians' consideration of joining.	No clear evidence in favor of alternative explanation that recruits were selected on the basis of their productive potential rather than being especially motivated to join on the basis of productive potential.
Table E.4	Relationship between own education and participation is not spurious to either household pre-war wealth or number of brothers.	No clear evidence in favor of alternative explanation that educational attainment is a reflection of whether household was sufficiently well-off that it could "afford" to have some sons join the rebellion.
Table E.5	Relationship between own education and participation is not spurious to measures of collective action resources, such as radio ownership, having been informed by a rebel group about joining, or tending to spend time with one's peers.	No clear evidence in favor of alternative explanation that educational attainment allowed acquisition of personal resources to facilitate participation in collective action.
Table E.6	Relationship between own education and participation is not spurious to levels of pre-war political support for different parties.	No clear evidence in favor of alternative explanation that educational attainment caused political socialization or radicalization that led to insurgent participation.

F Statistical robustness checks

F.1 Robustness to out-migration likelihood

Table F.1: Participation in revolt as a function of own and father's education

	1. Full sample	2. No border prov.	3. Full sample	4. No border prov.
(Constant)	0.051^{*}	0.030	0.054^{**}	0.041^{*}
	(0.020)	(0.021)	(0.020)	(0.017)
Age in 1972	-0.001^{*}	-0.001	-0.001^{*}	-0.001
	(0.000)	(0.001)	(0.000)	(0.001)
Age in 1972 sq.	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Subject's ed. (yrs.)	0.003^{***}	$0.002^{\dagger}$		
	(0.001)	(0.001)		
Logit survival pr.	-0.007^{**}	-0.009^{**}	-0.006^{**}	-0.009^{**}
	(0.002)	(0.003)	(0.002)	(0.003)
Father's ed. (yrs.)			0.003^{*}	0.002
			(0.001)	(0.002)
N	905	424	905	424
R^2	0.025	0.024	0.022	0.024

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

Dependent variable is participation in armed revolt (0,1).

All models control for pre-war province fixed effects.

Models 2 and 4 drop observations in border provinces.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table F.1 presents a robustness check that attempts to account for the possibility that our estimates are biased or misleading due to the potential for outmigration during the war. The sample is divided roughly evenly between those who, just prior to the start of the war, resided in areas bordering other countries and those who resided in interior provinces that did not border any other country. Those residing in border provinces would have had an easier time out-migrating than those residing in interior provinces. As such, to the extent that out-migration biases our analysis, we would expect the results to be driven to a large extent by patterns in these border provinces. We would get closer to the "true" relationship between education and participation by limiting our analysis to those who would not have easily out-migrated. We do this by re-estimating the coefficients on the subject's own and the subject's father's education using just the subsample of subjects who resided in interior provinces. The results are shown in Table F.1. As a basis for comparison, Models 1 and 3 replicate the full-sample results from Models 5 and 6 of Table 2. Then, Models 2 and 4 estimate the same quantities, but on the subsample of individuals whose location at just prior to the war was in an interior province. We see that the estimated coefficients are nearly identical. The sample is cut in half in this robustness check, and so it is no surprise that the statistical

significance level drops. But there is nothing here to suggest that our basic results are biased or misleading from aggregating over places that differ in terms of out-migration opportunities.

F.2 Logistic regression results

Table F.2: Participation in revolt as a function of own and father's education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Participate in rebellion (0/1)									
Age	-0.058**	-0.040*	-0.046*	-0.038*	-0.074***	-0.053**	-0.067***	-0.049**	-0.054***
	(0.018)	(0.016)	(0.018)	(0.016)	(0.020)	(0.018)	(0.019)	(0.018)	(0.015)
Age squared	0.002	0.001	0.001	0.001	0.003	0.002	0.003		
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		
Subject's ed. (yrs.)	0.141***		0.197***		0.165***		0.156***		
	(0.037)		(0.032)		(0.033)		(0.032)		
Father's ed. (yrs.)		0.095**		0.074***		0.081**	0.054	0.075*	0.071*
		(0.033)		(0.021)		(0.031)	(0.031)	(0.031)	(0.031)
Logit survival pr.					-0.345***	-0.306***	-0.329***	-0.301***	-0.307***
					(0.083)	(0.084)	(0.085)	(0.084)	(0.085)
Age in 1972 > 12								-0.076	
								(0.481)	
Father's ed. (yrs.) X Age in 1972 > 12								0.199	
								(0.145)	
Age in 1972 > 16									-0.033
									(0.637)
Father's ed. (yrs.) X Age in 1972 > 16									0.588***
									(0.165)
Observations	903	903	903	903	903	903	903	903	903

Standard errors in parentheses
Logistic regression coefficients reported.
Models 1-2 and 5-9 are weighted logistic regression with pre-war province fixed effects.
(Weights are the same as those used in the main text.)
Models 3-4 are unweighted conditional logistic regressions grouping by pre-war province.
Standard errors account for commune-level clustering (number of clusters is 65) except for models 3-4, which account for grouping factor clustering.
Models 2 and 4 use age-specific fixed effects.
$^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

F.3 Results without the age restriction on the sample

The results in the main text worked with a restricted sample that included only Hutu males who were between 12 and 42 years of age in 1993. The lower cutoff was based on the idea that it would only be for males of at 12 years of age that the arguments regarding education would be relevant (younger youth would not have had the chance to attain substantial years of education). The upper cutoff was based on the fact that the oldest rebel in our sample was 42 in 1993, and so the insurgency participation rate for those above 42 years is exactly 0 in our sample. In the tables that follow, we replicate the statistical results from the paper without imposing these age restrictions. In qualitative terms, most of the estimates are the same in the restricted and unrestricted sample. In the unrestricted sample, some of the key estimates are attenuated (e.g., the estimated effect of father's education on participation) and some no longer hold (e.g., the interaction coefficient in the cohort-based analysis), which is to be expected given that the unrestricted sample is includes units that provide little relevant variation, and in particular, because of the fact that among those older than 42, the fact that we have a large number of no-participation outcomes induces a non-linearity in the age-based conditional participation probability.

Table F.3: Participation in revolt as a function of own and father's education (full sample, without age restrictions)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
(Constant)	0.074^{***}	0.075^{***}	0.084^{***}	0.064^{**}	0.077^{***}	0.088^{***}	0.073^{***}
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Age in 1972	-0.001^{***}		-0.000		-0.002^{***}	-0.001^{***}	-0.002^{***}
	(0.000)		(0.000)		(0.000)	(0.000)	(0.000)
Age in 1972 sq.	-0.000		-0.000^{**}		0.000^{***}	0.000	0.000^{**}
	(0.000)		(0.000)		(0.000)	(0.000)	(0.000)
Subject's ed. (yrs.)	0.003^{***}	0.003^{***}			0.003^{***}		0.003^{***}
	(0.001)	(0.001)			(0.001)		(0.001)
Father's ed. (yrs.)			0.002^{**}	0.002^{**}		0.002^{*}	$0.001^{\dagger}$
			(0.001)	(0.001)		(0.001)	(0.001)
Logit survival pr.					-0.007^{***}	-0.007^{***}	-0.007^{***}
					(0.002)	(0.001)	(0.001)
N	1460	1460	1460	1460	1460	1460	1460
R^2	0.021	0.027	0.019	0.027	0.030	0.027	0.031
adj. R^2	0.006	-0.015	0.004	-0.016	0.015	0.012	0.016
Resid. sd	4.298	4.344	4.303	4.346	4.278	4.285	4.278

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

(Number of clusters is 65.)

Dependent variable is participation in armed revolt (0,1).

All models control for pre-war province fixed effects.

Models 2 and 4 use age-specific fixed effects.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table F.4: Subject's education and occupational ranking as a function of father's education (all-ethnicity sample; full sample, without age restrictions))

	Pre-war ed.	Pre-war occ. ranking
(Constant)	6.048^{***}	0.129
	(0.582)	(0.115)
Age in 1972	0.180^{***}	0.043^{***}
	(0.034)	(0.006)
Age in 1972 sq.	-0.014^{***}	-0.002^{***}
	(0.003)	(0.001)
Father's ed. (yrs.)	0.175^{***}	-0.010
	(0.045)	(0.010)
Hutu	$-1.087^{\dagger}$	-0.048
	(0.591)	(0.098)
Father's ed.X Hutu	-0.064	0.021
	(0.058)	(0.014)
Age X Hutu	-0.082^{*}	-0.005
	(0.034)	(0.006)
Age sq. X Hutu	0.006	0.000
	(0.003)	(0.001)
N	1779	1779
R^2	0.375	0.413
adj. R^2	0.366	0.405
Resid. sd	86.126	17.661

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

(Number of clusters is 66.)

D.V. in Model 1 is respondent's own level of education before the war (0-18).

D.V. in Model 2 is respondent's own occupational ranking before the war (-1,0,1).

Sample includes both Hutu and Tutsi male respondents.

All models control for pre-war province fixed effects and survival probability.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table F.5: Occupational ranking as a function of own and father's education

	Pre-war occ.	Pre-war occ.	Post-war occ.	Post-war occ.
(Constant)	-0.246^{*}	0.015	0.078	0.257^{**}
	(0.096)	(0.099)	(0.089)	(0.094)
Age in 1972	0.032^{***}	0.039^{***}	-0.012^{**}	$-0.007^{\dagger}$
	(0.003)	(0.003)	(0.004)	(0.004)
Age in 1972 sq.	-0.001^{**}	-0.002^{***}	$0.000^{\dagger}$	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Subject's Ed. (yrs.)	0.060^{***}		0.042^{***}	
	(0.010)		(0.007)	
Father's ed. (yrs.)		0.010		0.008
		(0.009)		(0.007)
N	1460	1460	1460	1460
R^2	0.431	0.381	0.155	0.113
adj. R^2	0.423	0.372	0.142	0.100
Resid. sd	16.188	16.887	14.945	15.311

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

(Number of clusters is 65.)

D.V. is occupational rank (-1,0,1).

All models control for pre-war province fixed effects and survival probability.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

[1] 66 [1] 66

Table F.6: Military ranks as a function of own and father's education (rebels only)

	Incoming rank	Incoming rank	Outgoing rank	Outgoing rank
(Constant)	0.281	1.953***	1.152**	6.136***
	(0.347)	(0.282)	(0.431)	(0.525)
Age in 1972	−0.014	0.005	−0.062**	−0.028
	(0.019)	(0.020)	(0.020)	(0.022)
Age in 1972 sq.	0.000	−0.001	0.001	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Subject's Ed. (yrs.)	0.219***		0.642***	
	(0.029)		(0.042)	
Father's ed. (yrs.)		0.037		0.195***
		(0.024)		(0.033)
N	704	704	706	706
R^2	0.235	0.165	0.419	0.221
adj. R^2	0.210	0.138	0.400	0.196
Resid. sd	9.495	9.918	13.578	15.721

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

(Number of clusters is 65.)

D.V. is military rank (1 simple soldier - 15 general).

All models control for pre-war province fixed effects and survival probability.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table F.7: Education as a function of father's education, by age categories in 1972

	Model 1	Model 2	Model 3
(Constant)	4.532^{***}	4.696^{***}	4.754^{***}
	(0.311)	(0.366)	(0.290)
Age in 1972	0.111^{***}	0.074^{*}	0.081^{***}
	(0.020)	(0.029)	(0.022)
Age in 1972 sq.	-0.009^{***}	-0.013^{***}	-0.013^{***}
	(0.002)	(0.002)	(0.002)
Father's ed. (yrs.)	0.117^{**}	0.081^{*}	0.090^{*}
	(0.042)	(0.037)	(0.037)
Age in 1972 > 12		1.611	
		(1.033)	
Father's ed. (yrs.) X Age in 1972 > 12		0.821^{***}	
		(0.147)	
Age in 1972 > 16			2.256^{*}
			(1.065)
Father's ed. (yrs.) X Age in 1972 > 16			0.641^{***}
			(0.117)
N	1460	1460	1460
R^2	0.301	0.330	0.322
adj. R^2	0.290	0.319	0.311
Resid. sd	80.601	78.951	79.405

Weighted least squares with standard errors accounting for clustering at the commune level.

Number of clusters is 65.

Dependent variable is respondent's years of education before the war (0-18).

All models control for pre-war province fixed effects and survival probability.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table F.8: Participation in revolt as a function of father's education, by age categories in 1972 (full sample, without age restrictions)

	Model 1	Model 2	Model 3
(Constant)	0.088^{***}	0.087^{***}	0.088^{***}
	(0.022)	(0.022)	(0.022)
Age in 1972	-0.001^{***}	-0.001^{***}	-0.001^{***}
	(0.000)	(0.000)	(0.000)
Age in 1972 sq.	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Father's ed. (yrs.)	0.002^{*}	0.002^{*}	0.002^{*}
	(0.001)	(0.001)	(0.001)
Age in 1972 > 12		$0.016^{\dagger}$	
		(0.009)	
Father's ed. (yrs.) X Age in 1972 > 12		-0.001	
		(0.003)	
Age in 1972 > 16			0.023^{*}
			(0.010)
Father's ed. (yrs.) X Age in 1972 > 16			0.000
			(0.004)
N	1460	1460	1460
R^2	0.027	0.028	0.028
adj. R^2	0.012	0.011	0.011
Resid. sd	4.285	4.287	4.287

Weighted least squares estimates.

Standard errors account for clustering at the commune level.

Number of clusters is 65.

Dependent variable is participation in armed revolt (0,1).

All models control for pre-war province fixed effects and survival probability.

† significant at $p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

‡ jointly significant at $p < .10$.

References

Angrist, Joshua D. 1997. "Conditional Independence in Sample Selection Models." *Economics Letters* 54(2):103–112.

Atran, Scott. 2010. *Taking to the Enemy: Violent Extremism, Sacred Values, and What it Means to Be Human*. New York, NY York: Penguin.

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

Blattman, Christopher. 2009. "From Violence to Voting: War and Political Participation in Uganda." *American Political Science Review* 103(2):231–247.

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

Chinodya, Shimmer. 1990. *Harvest of Thorns*. London: Heinemann.

Clarke, Kevin A. and David M. Primo. 2012. *A Model Discipline: Political Science and the Logic of Representations*. Oxford: Oxford University Press.

Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil Wars." *Oxford Economic Papers* 56(4):563–594.

de Tocqueville, Alexis. 1856. *The Old Regime and the French Revolution*. NY: Harper and Bros.

DeNardo, James. 1985. *Power in Numbers: The Political Strategy of Protest and Rebellion*. Princeton, NJ: Princeton University Press.

Esteban, Joan and Debraj Ray. 2011. "A Model of Ethnic Conflict." *Journal of the European Economic Association* 9(3):496–521.

Esteban, Joan, Laura Mayoral and Debraj Ray. 2012. "Ethnicity and Conflict: An Empirical Study." *American Economic Review* 102(4):1310–1342.

Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau and Yu-Sung Su. 2008. "A Weakly Informative Default Prior for Logistic and Other Regression Models." *Annals of Applied Statistics* 2(4):1360–1383.

Gilligan, Michael, Eric Mvukiyehe and Cyrus Samii. 2013. "Reintegrating Rebels into Civilian Life: Quasi-Experimental Evidence from Burundi." *Journal of Conflict Resolution* 57(4):(In press).

Glaeser, Edward L., Giacomo A. M. Ponzetto and Andrei Shleifer. 2007. "Why Does Democracy Need Education?" *Journal of Economic Growth* 12(2):77–99.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83(6):1420–1443.

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

Press.

Humphreys, Macartan and Jeremy M. Weinstein. 2008. "Who Fights? The Determinants of Participation in Civil War." *American Journal of Political Science* 53(2):436–455.

Jackson, Tony. 2000. *Equal Access to Education: A Peace Imperative for Burundi*. London: International Alert.

Kalyvas, Stathis and Matthew Adam Kocher. 2007. "How "Free" is Free Riding in Civil Wars?: Violence, Insurgency, and the Collective Action Problem." *World Politics* 59(2):177–216.

Korn, Edward L. and Barry I. Graubard. 1999. *Analysis of Health Surveys*. New York, NY: Wiley.

Krueger, Alan B. 2007. *What Makes a Terrorist: Economics and Roots of Terrorism*. Princeton, NJ: Princeton University Press.

Kuran, Timur. 1989. "Sparks and Prarie fires: A theory of unanticipated political revolution." *Public Choice* 61:41–74.

Lee, Alexander. 2011. "Who Becomes a Terrorist? Poverty, Education, and the Origins of Political Violence." *World Politics* 63(2):203–245.

Lemarchand, Rene. 1994. *Burundi: Ethnic Conflict and Genocide*. Cambridge University Press.

Manirakiza, Marc. 1992. *De la révolution au régionalisme: Burundi, 1966-1976*. Brussels: La Longue Vue.

Manski, Charles F. 1995. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.

Medina, Luis Fernando. 2007. *A Unified Theory of Collective Action and Social Change*. University of Chicago Press.

Medina, Luis Fernando. 2013. "The Analytical Foundations of Collective Action Theory: A Survey of Some Recent Developments." *Annual Review of Political Science* 16(1):259–283.

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

Newey, Whitney K. 2009. "Two-Step Series Estimation of Sample Selection Models." *The Econometrics Journal* 12(s1):S217–S229.

Ngaruko, Floribert and Janvier D. Nkurunziza. 2000. "An Economic Interpretation of the Conflict in Burundi." *Journal of African Economies* 9(3):370–409.

Olson, James M. and Neal J. Roese. 2002. Relative Deprivation and Counterfactual Thinking. In *Relative Deprivation: Specification, Development, and Integration*, ed. Iain Walker and Heather J. Smith. Cambridge: Cambridge University Press.

Reyntjens, Filip. 1994. *L'afrique des grands lacs en crise*. Paris: Karthala.

Schelling, Thomas. 1978. *Micromotives and Macrobehavior*. New York, NY: W. W. Norton.

United Nations. 1996. "United Nations International Commission of Inquiry for Burundi: Final

Report." New York.

VanBelle, Douglas A. 1996. "Leadership and Collective Action: The Case of Revolution." *International Studies Quarterly* 40:107–132.

Wood, Elisabeth Jean. 2003. *Insurgent Collective Action and Civil War in El Salvador*. Cambridge University Press.