WARWICK

ECONOMICS

Mistaking Noise for Bias Victimhood and Hutu-Tutsi Reconciliation in East Africa

Arthur Blouin & Sharun W. Mukand

This paper is also included in the CAGE Discussion paper series

July 2022 No: 1417

Warwick Economics Research Papers

ISSN 2059-4283 (online)

ISSN 0083-7350 (print)

Mistaking Noise for Bias

Victimhood and Hutu-Tutsi Reconciliation in East Africa

Arthur Blouin
University of Toronto

Sharun W. Mukand
University of Warwick

June 24, 2022*

Abstract: The difficulty in resurrecting inter-ethnic cooperation in the aftermath of violence and genocide is one of the biggest challenges facing post-conflict societies. Using experimental data from post-genocide Rwanda and Burundi, this paper shows that an unwarranted tendency to blame others for negative outcomes is a behavioural barrier that makes reconciliation difficult. We show that individuals systematically (and mistakenly) blame accidental negative shocks (noise) to the deliberate intent of individuals (bias). This "victimhood bias" wherein individuals ascribe noise to bias is much larger for (a) individuals for whom ethnic identity is salient; (b) for those who have had greater exposure to inter-ethnic violence. Further, we observe that *both* inter-ethnic contact *and* economic development are associated with a decline in this victimhood bias. Finally, those with a lower victimhood bias are more likely to behave cooperatively in inter-ethnic relationships. Our results suggest that insurance agreements that limit negative shocks and reduce noise, can encourage reconciliation by mitigating feelings of victimhood.

*We are extremely grateful for financial support from the British Academy and the Centre for Competitive Advantage in the Global Economy at the University of Warwick. We would also like to acknowledge the tremendous support of our field teams in Rwanda and Burundi, including several individuals who offered invaluable insight into aspects of Hutu/Tutsi culture. Without their help the project is unlikely to have succeeded. We refrain from acknowledging any of them individually (they still live in the region) to protect their anonymity.

1. Introduction

Over a third of all nations have experienced civil conflict during the past half century (Blattman and Miguel, 2010). The shadow of past and ongoing conflict continues to hang over much of the developing world and has contributed to ethnic polarization and mistrust, weak institutions and disappointing economic growth (Easterly and Levine, 1997, Garcia-Montalvo and Reynal-Querol, 2004, Michalopoulos and Papaioannou, 2016). Though it has proven difficult to achieve, the restoration of trust and the promotion of inter-ethnic cooperation across society (Arrow, 1970, 1972) has the potential to promote politically sustainable growth and stability. In this paper, we re-examine the behavioural roots of inter-ethnic reconciliation to ask: why is it so difficult to resurrect inter-ethnic cooperation in the aftermath of conflict?

It should not surprise us that cooperation in fractured relationships is hard to repair (e.g. Fehr and Gächter (2000), Gächter et al. (2017)) since they can result in mutual suspicion. Indeed, it is rational (and Bayesian) for individuals to react with caution in the face of a negative shock. This is especially the case because in any relationship, if there is a negative shock, it may be difficult to distinguish "noise" (i.e. accidental random shocks) from "bias" (i.e. correlated shocks) arising from the deliberate and prejudicial actions of others.[1] However, the sustainability of all relationships is predicated on the ability of information to be processed in an unbiased manner (Holmström, 1979). If accidental negative shocks result in *unwarranted* blame on the other party, then cooperation can be much harder to achieve.

We explore whether individuals systematically mistake noise for bias using data from a series of lab-in-the-field experiments with 774 subjects in 136 villages across Burundi and Rwanda.[2] This region has witnessed violent ethnic conflict for several decades, including one of the worst genocides in recent history in 1994, when over 70 percent of the Tutsi minority in Rwanda was massacred. This context provides an apt laboratory to examine trust, misunderstanding, and cooperation in inter-ethnic relationships.

The key experimental intervention that we introduce is the *victim game* - a

[1] Of course, given our inter-ethnic context we are playing on the term "bias" to not only designate any correlated errors in judgement, but more specifically *apply to errors in judgement that arise due to prejudice* on the part of one of the individuals in the (possibly inter-ethnic) relationship. See Kahneman et al. (2021) for a detailed discussion on the statistical definition and distinction between noise and bias.

[2] Hutus and Tutsis constitute an overwhelming majority of the population in Burundi and Rwanda.

stochastic version of the trust game (see Kosfeld et al. (2005)).[3] In this game, just as in the trust game, one player (the sender) decides to share some fraction of an endowment with the other player (the receiver). Whatever is shared is doubled, and then the receiver decides how to divide that money between the two players. However, with some probability the amount originally shared by the sender is determined randomly, so the receiver faces uncertainty over whether the money they receive is from the sender or is randomly selected by a roll of the dice (i.e. "noise"). The receiver observes the identity of the possible sender as well as the amount of money received. They are then asked to ascribe the source of the income received either to luck (i.e. roll of the dice) or to the deliberate choice of the sender. Analysing the data from this victim game, in conjunction with other ancillary interventions, generates a rich set of findings.

Our first set of results report a behavioural barrier that can constrain reconciliation and make it difficult to restore cooperation. We show that individuals in economic relationships suffer from a systematic confusion, insofar that they mistakenly ascribe accidental shocks to the deliberate intent of the other party - especially when the shocks are negative.[4] Moreover, this *victimhood bias*, which results in mis-ascription of noise to bias, is particularly severe when the relationship is inter-ethnic. Since the effect is particularly strong both when shocks are negative and when partnerships are inter-ethnic, this suggests that subjects in these post inter-ethnic conflict regions are quicker to adopt a sense of victimhood. The severity of this victimhood bias in inter-ethnic relationships is particularly surprising, since it is unwarranted given actual behaviour that we observe in both Rwanda and Burundi. In particular, in both these countries individuals treat those who belong to the other party in a very similar manner with respect to the monetary offers, whether or not they are in a co-ethnic or an inter-ethnic relationship.

We further examine the underpinnings of this *victimhood bias*. We adapted Taylor et al. (1978) and Kurzban et al. (2001) to develop a measure of Hutu and Tutsi ethnic salience (see Blouin and Mukand (2019)). We use this measure to

[3] A randomization device has been introduced to variants of the trust game by Kosfeld et al. (2005) as well as Bohnet and Zeckhauser (2004) who use it to study betrayal aversion.

[4] That is, they are quick to (incorrectly) blame. There is a literature on blame in both Economics (e.g. Charness and Levine (2007), Falk et al. (2008), Gurdal et al. (2013)) and Psychology (e.g. Mazzocco et al. (2004), Gino et al. (2008)). However, the aim of this literature is quite different from ours. The literature on blame typically seeks to explain *why* a person might sometimes blame others times not. Instead, we are interested in explaining *who* is most likely wrongful blames whom in the field - to specifically explore how blame relates to conflict and reconciliation.

show that among subjects for whom ethnicity is particularly salient, there is a greater tendency to blame negative economic outcomes onto those of the other ethnicity. Furthermore, the victimhood bias in inter-ethnic relationships is much more severe in regions where there was a history of direct exposure to inter-ethnic violence.

Finally, we offer some suggestive evidence regarding the benefits of social or institutional arrangements that can mitigate the issue for inter-ethnic relationships. For instance, we observe that individuals who were born in ethnically diverse villages - and prospectively experienced greater inter-ethnic contact - are *less* likely to suffer from a victimhood bias. Consistent with this, subjects who are more likely to blame members of the other ethnic group for bad outcomes are less likely to choose to partner with out-group members in a co-operative task. Furthermore, our results show that economic development is correlated with a reduction in this victimhood bias that may plague inter-ethnic relationships. Taken together, these results suggest that policies that mitigate negative economic shocks - like public insurance policies - would be especially important to consider in post-conflict contexts, and could be instrumental to promoting post-conflict reconciliation.

This paper contributes to the literature on ethnic reconciliation in conflict societies (Paluck et al., 2021). The majority of work on reconciliation has focused on improving trust (Annan et al., 2011, Bauer et al., 2014, Bellows and Miguel, 2009, Cassar et al., 2013, Rohner et al., 2013). In our context cooperation remains low despite high trust, highlighting that trust is not the only barrier to reconciliation. This can have important policy implications, since changing attitudes through the media has been pillar of reconciliation policy around the world (Bratic, 2013, DellaVigna et al., 2014, Rao, 2014, Armand et al., 2020). For instance, in our regional context, Paluck and Green (2009), Paluck (2012) and Blouin and Mukand (2019) examine how exposure to the radio promoted reconciliation in Rwanda. Our findings suggest that policies that mitigate negative economic shocks might be another important policy tool in addition to the current media-based strategies, further reinforcing the difficulty and nuance in fostering reconciliation in post-conflict societies (Shnabel and Nadler, 2008, 2015).

Our work is also related to the work in social psychology on the fundamental attribution error (Ross, 1977, Gilbert, 1998) in the context of inter group relationships (Pettigrew, 1979). The fundamental attribution error examines the tendency to attribute the intent of someone's action to their personality rather than situational factors. In other words, the focus is on *why* an individual chose a particular action, whereas we focus on perceptions of *whether* an observed outcome was due

to chance or individual behaviour in the first place.

Finally, our work offers a very different perspective on recent work on noise by (Kahneman et al., 2021). While this work emphasises the importance of "decision hygiene" in the organisational context, our findings suggest that noise can have important implications in very different social and institutional contexts - including the study of conflict resolution and ethnic reconciliation. At the same time, our results suggest important differences. For instance, they argue that "bias and noise play the same role in the calculation of overall error...*a reduction of noise has the same impact on overall error as does a reduction in bias by the same amount*" (Kahneman et al., 2021, p.55, emphasis added). While they emphasize the independence of noise and bias, in social settings these may not be independent. We show that noisy decisions (i.e. randomly low inter-ethnic offers) can be systematically confused for bias, leading to inefficient responses that *increase* the overall error. This suggests that the broader message of Kahneman et al. (2021) applies with even more emphasis - since a reduction in noise will have not equal, but a disproportionate impact by also reducing the *perception* of bias.

2. Data and Empirical Strategy

Our data comes from lab-in-the-field experiments and associated surveys conducted in rural Rwanda using 438 farmers from 52 different villages and an additional 329 farmers from 82 villages in rural Burundi (figure B1). We chose villages on the basis of the geographic suitability of the land for coffee production,[5] and restricted our selection to 'FARG-eligible' villages in Rwanda. In Rwanda it is not permissible to directly ask ethnicity, and FARG-eligibility helps to distinguish between Hutu and Tutsi in the sample.[6] Given this selection, we ended up with just under 30% of our sample belonging to the minority Tutsi in Rwanda. In contrast, not only were we able to directly identify and ask ethnicity in Burundi, the Tutsi were also more evenly geographically distributed and Tutsi also constitute just under 30% of our Burundi sample.

[5]The reason is that this is a useful proxy for whether the village had a history of forced labour - please see: Blouin (2022) for details.

[6]In Rwanda under Paul Kagame, it is not legally permissible to directly ask individuals their ethnicity. However, as described in Blouin and Mukand (2019) we discovered (and subsequently validated) a standard survey question in Rwanda, on sources of government assistance received by each farmer, one source of funding was a very good proxy for ethnicity. In particular, we discovered that almost all Tutsis received funds from the government from the FARG ("Fonds d'Assistance pour Rescapées du Génocide") program that is exclusively targeted to Tutsi families who are 'survivors' of the 1994 genocide. Accordingly, an affirmative response to this question, was a very good proxy for Tutsi ethnicity.

In each village that was chosen for a survey and experiment, we chose participants randomly from a list of subjects made available from each village. We compensated all those who agreed to participate in the experiment well by providing them with a case of soap (more than a day's wage).[7] Not surprisingly, in the vast majority of villages everybody participated. We ensured that all individuals who participated in the various experiments were randomly matched to play with individuals from a different village and had never met each other. The total time taken for a respondent to complete the experiments as well as the survey responses was less than half a day. Summary statistics for our main outcomes and controls can be seen in table B1.

Below, we briefly describe our experimental measures and protocol for Rwanda and relegate details to Appendix A as well as Blouin and Mukand (2019).

2.A. The victim game

The victim game is a simple variation of the standard trust game that is used to elicit inter-group attitudes (Fershtman and Gneezy, 2001). As in the trust game, a randomly matched pair of individuals were selected to play face-to-face experiments with the help of an enumerator, who was conducting the experiment. Also as in the trust game, one individual was randomly designated to be the 'sender' and the other the 'receiver'. Each subject in the pairing was selected from different villages and prior to the game an enumerator confirmed that they had never previously met. Moreover, in order to minimize strategic considerations, each pair was only allowed to play one round.

At the beginning of the game the enumerator gave the 'sender' 600RWF (approximately one US dollar). We should emphasize that in this (very poor) region, this amount is a substantial amount of money and introduces real stakes for all subjects – since it nearly equals the average earnings from a hard day of physical labour. The sender could share some (discrete) amount x of this money with the 'receiver' (i.e. their partner) and keep the remainder $600 - x$ for themselves, where $x \in \{0, 100, 200, ..., 600\}$. The enumerator matched the amount shared by the sender. At this stage the victim game diverges from the standard trust game in one important respect.

In particular, both the sender and receiver were informed that the enumerator would carry out a dice roll, whose outcome would *not* be revealed to them. However, we did inform them that if the dice roll was a four, five or six, the receiver

[7] Often people in the rural areas we ampled have to travel into the city to buy soap, so soap as compensation brings additional benefits of saving up to a day of travel.

obtained the amount $2x$, where as in the standard trust game, x was chosen by the sender. In contrast, if the dice roll was a one, two or three they received a randomly determined amount that was left unspecified. In practice, if the dice role was a one, they received twice 100RWF; if it was a two they receive two times 300RWF; and if it was a three they receive two times 500RWF. From the money they received, the receiver returned some amount y to the sender. Payoffs were then determined with the sender getting $(600 - x) + y$ and the receiver's payoff equaling $2x - y$. Importantly, the players were *not* informed whether the source of the money received by the receiver, was the sender or the dice-roll.

Once the regular game payoffs were realized, the receiver had an additional opportunity to earn money. They were asked to reveal whether they believe the offer they received came from the dice-roll, or from their partner. If the receiver guessed correctly, and attributed the money received (i.e. $2x$) to the correct source - be it the sender or the random roll of the dice - (s)he received a 200RWF bonus.

2.B. Salience of Ethnic Identity:

We are also interested in assessing whether any victimhood bias that we observe is correlated with whether subjects 'categorize' others on the basis of ethnicity.[8] Accordingly, we adapt the experimental design developed in Taylor et al. (1978) and Taylor and Fiske (1978) to our Hutu-Tutsi context.[9]

The complete details of our experimental protocol are described in Blouin and Mukand (2019). In broad terms, the exercise entailed a memory recall task, where individuals may mistakenly attribute statements made by one individual to another individual. If a subject was systematically more likely to misattribute a statement associated with a Tutsi to another Tutsi, rather than a Hutu, then we would say that for the subject, ethnicity was more likely salient. Accordingly, we interpret a greater share of errors that are within-ethnicity as a measure of ethnic salience. This is because a subject may make fewer errors due to Taylor et al. (1978) observation that an individual "may select salient social or physical

[8]Any such categorization presumes that there are discernible physical/genetic differences between the Hutu and the Tutsi. This is discussed at some length in Blouin and Mukand (2019). The upshot of the literature is that while there is a clear element of political construction of identity in that these ethnic categories have 'hardened' during colonization. However, there are genetic differences (on average) between Hutu and Tutsi as demonstrated by several genetic studies including (Luis et al. (2004) and Shepard and Herrera (2006)).

[9]This is often known as the "memory confusion protocol" and the original Taylor et al. (1978) experiment was used to study whether individuals encode race. The importance of such categorization for social cognition has been further explored by Stangor et al. (1992) and Kurzban et al. (2001).

dimensions...for grouping and managing personal information." As described in Blouin and Mukand (2019), our measure of the salience of identity is given by the following:

$$\text{ethnic salience} = \frac{\sum \text{within ethnicity errors}}{\sum \text{errors}}.$$

An attractive aspect of this measure of ethnicity salience is that it is unobtrusive and captures in a simple way how a subject processes and categorizes information about others identity without priming them about their ethnicity. This is imperative in Rwanda where even the mention of Hutu or Tutsi is strongly discouraged by the authoritarian goverenment.

2.C. Empirical Approach

The main hypothesis that we aim to test is that in Rwanda and Burundi - regions of the world that have experienced immense ethnic conflict - individuals tend to mistake random negative shocks as intentional, and are especially likely to do so in inter-ethnic contexts. Towards that end, we rely on two randomized lab-based interventions. We then explore heterogeneity in this core result by examining different regions, and different individual characteristics.

As mentioned earlier, individuals were randomly allocated partners in the victim game. This implies that whether a person was partnered with someone of their own ethnic group or another is also random. Recollect that we are interested in an individual's proclivity to mistakenly attribute to deliberate human intent what are essentially random negative or positive shocks. Accordingly, our empirical analysis focuses on the random offers; when the enumerator rolls a one, two, or three. Our focus on how subjects respond to randomly generated high and low offers also has an important additional advantage: in particular, it allows us to directly use the human offers to control for priors, proxied for with the distribution of actual offers.[10,11]

[10] If the human offers were included in the main analysis, we would have both the issue of non-random high and low offers to address - which serves as our main treatment - but we would also have some function of human offers on the left and right hand sides of the regression equation.

[11] Instead of relying on elicited priors, which would have required that we make ethnicity salient, we use the observed distribution of offers of observationally equivalent subjects. In particular, we want to control for the Bayesian, rational belief about how accountable is the partner upon receiving a particular offer. We define a prior $P(\text{human}|\text{offer})$ to be the share of people from the session that are observationally equivalent to the respondent. We use observational equivalence based on the following criteria: age band (5-year bands), gender, ethnicity and partner ethnicity. that offered the amount that the respondent received. We then compute: $P(\text{offer}|\text{human}) = \frac{P(\text{human}) \cdot P(\text{human}|\text{offer})}{P(\text{offer})}$

Accordingly, our key specification is,

$$(1) \qquad Victim_{is} = \beta_0 + \beta_1 InterEthnic_{is} \cdot LowOffer_{is} + \beta_2 InterEthnic_{is} + \beta_3 LowOffer_{is} + X'_{is}\Gamma + \epsilon_{is}$$

Where i denotes an individual and s denotes the lab-session that they attended. Random offers can be 100RWF, 300RWF or 500RWF, and here we denote $LowOffer_{is}$ to be the offers of 100RWF. $Victim$ is binary and equal to 1 if the individual attributes the random offer to their partner, so β_1 in this case can be interpreted as the differential blame directed towards the human partner when an offer is low, for an inter-ethnic partnership.

We focus on low offers because of the interest in victimhood and reconciliation. It facilitates interpretation to focus on the low offers in the regression, however, throughout the paper we present graphically the estimated distribution for the full range of offers. X'_{is} is a vector of controls, it includes gender, ethnicity, age, a measure of IQ, Province dummies and a variable capturing priors.

Once we estimate β_1, we explore heterogeneity in the effect. In subsequent regressions, we limit the sample only to the inter-ethnic pairings, and focus on more suggestive results of the following type:

$$(2) \qquad Victim_{is} = \beta_0 + \beta_1 Characteristic_{is} \cdot LowOffer_{is} + \beta_2 Characteristic_{is} + \beta_3 LowOffer_{is} + X'_{is}\Gamma + \epsilon_{is}$$

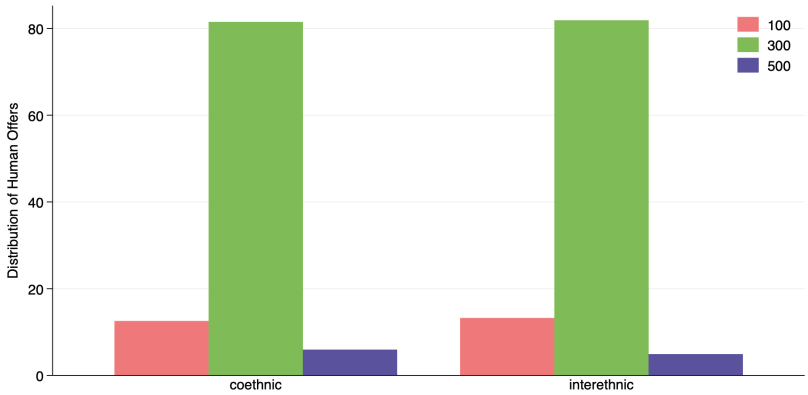
Where $Characteristic_{is}$ is a dimension of interest with respect to heterogeneity of the effect. For instance, we explore the respondents' ethnic salience and the history of ethnic violence amongst other factors. These treatment effect heterogeneity effects are not intended to necessarily be interpreted causally, but they highlight which factors are associated with the main effect that we focus on.

3. Results

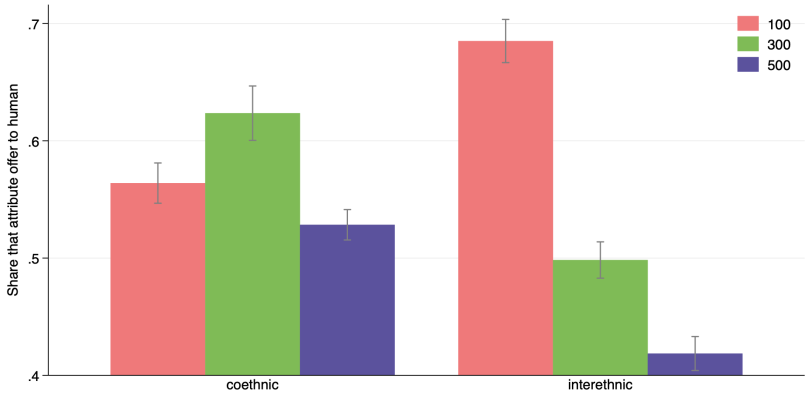
3.A. Mistaking Noise for Bias

In order to assess whether there is any systematic and erroneous attribution of offers by the receiver, we first describe the money offered by the sender in the victim game.

i) The Distribution of Sender Offers: Our results are depicted in figure 1a. Two aspects of the initial offers made by the sender are worth noting. First, an overwhelming majority of the offers made in the victim game are 300RWF with few subjects making low offers and even fewer making very high offers. Second, we observe that there is considerable similarity between co-ethnic and inter-ethnic offers made by the sender in the victim game. In what follows, we use these initial offers as a benchmark against which we can identify a tendency for misattribution.



(a) Sender offers in the victim game



(b) Ascription of offer to Sender (and not Luck)

Figure 1: The Victimhood Bias

Note: Figure (a) plots the distribution of offers in the victim game that were low, medium or high. In figure (b) we plot the rate at which individuals hold their partners accountable for the offer received. In both figures we plot the predicted values after regressing the y-axis variable on the controls and the interaction between inter-ethnic partnership and the offer observed. Controls include gender, ethnicity, age, age-squared, beliefs, IQ. The sample includes all subjects who received a random offer.

ii) The Distribution of Misascription: To begin with, recollect that in the victim game the receiver is not aware whether any money received is the result of random luck or the result of an intentional offer by the sender. We also observe that the

randomized offers are uniformly distributed. Given this, if information processing is unbiased, then (for any set of priors) the ascription distribution should be a weighted combination of the two distributions (i.e. the actual distribution of offers and the uniform distribution due to the dice roll). We graphically depict our results in figure 1b. The first thing to observe is that upon receiving both a low or a high offer, subjects are (as expected) much more likely to hold their partner accountable, rather than attribute the outcome to mere luck. In particular, even among high offers in the inter-ethnic partnerships where attribution to human offers is at its lowest, still 40% of people attribute the offers to their partners despite the fact that these offers are rarely made by humans.

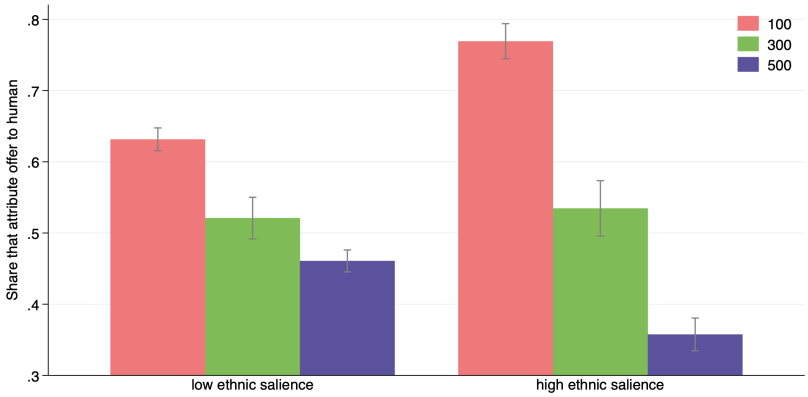
However, the most striking aspect of figure 1b is the dramatic difference between co-ethnic and inter-ethnic partnerships. In particular, subjects held their partners accountable for low inter-ethnic offers almost 70% of the time even though in practice, these offers occurred less than 20% of the time. They held their partners accountable for low offers nearly twice as often as for high offers, even though both occurred at a similar rate. Overall this suggests troubling ethnic patterns in the region - not with trust, surprisingly - but with people's tendency to blame their partners when things go wrong. Even with the considerable progress on reconciliation in the region, and the associated shocking progress on trust in inter-ethnic relations, progress on blame, and associated attitudes of victimization appear not to have progressed at the same rate.

3.B. Ethnic salience and violence:

Individuals have a tendency to blame others for low offers more than they should in inter-ethnic partnerships, but is this related to conflict and reconciliation? We can investigate the findings further by exploring heterogeneity in the results by ethnic salience, which has been linked with the reconciliation effort in Rwanda, as well as the conflict itself.

One key metric of reconciliation in the region is the extent to which people categorize others on the basis of ethnicity, and we find that on this dimension there is hope for recovery. We collected the salience of ethnic identity in both Rwanda and Burundi. Our measure is identical to that used in Blouin and Mukand (2019), and builds off the method developed in Taylor et al. (1978). Subjects saw eight photos of Hutu or Tutsi men (see figure 2a). For each photo, the enumerator picked up the photo, and read a neutral statement about the person in the photo, for instance, "this person likes to go for long walks." After reading each statement,

(a) Memory Confusion Protocol



(b) The role of ethnic salience

Figure 2: Victimhood Bias is more pronounced where salience remains high

Note: Figure (a) presents artist renders of the images used in the SIT exercise. Photos were used in the actual experiment, the artist altered facial characteristics to preserve the anonymity of the participants. In figure (b) we plot the rate at which individuals hold their partners accountable for the offer received. In both figures we plot the predicted values after regressing the y-axis variable on the controls and the interaction between ethnic salience and the offer observed. Ethnic salience is divided into high / low based on the median level of ethnic salience. Controls include gender, ethnicity, age, age-squared, beliefs, IQ. The sample includes all subjects who were assigned to a partner not from their own ethnic group, and who received a random offer.

subjects were asked to match a statement that was read back to them with the correct photograph. If they systematically confused a Hutu for another Hutu, or Tutsi for another Tutsi, we say that ethnic salience is high for them.

Figure 2b plots the tendency of people to hold their inter-ethnic partners accountable for low, medium or high offers, by high and low ethnic salience. For those with both high and low ethnic salience, we again see a large left-tail of individuals who wrongly blame their partners for bad offers, just as in the overall inter-ethnic sample. However, this pattern is much more prominent for the people with high ethnic salience than it is for those with low ethnic salience. This

seems intuitive, the people who are less likely to categorize others on the basis of ethnicity behave more similarly to how they would if they were in a co-ethnic pairing.

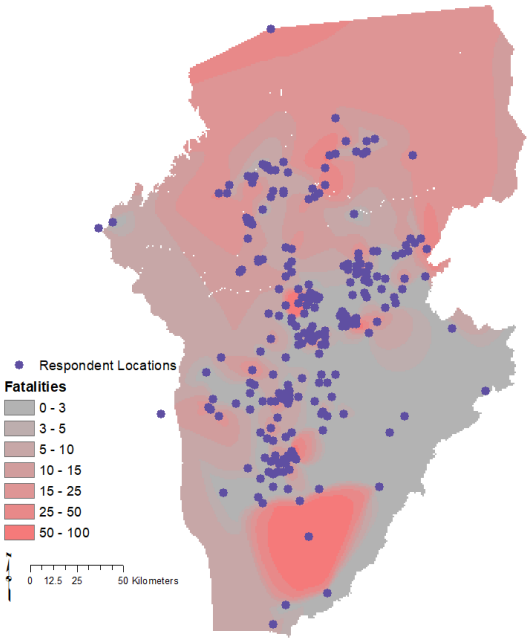
In the same vein, we can examine historical ethnic conflict. The geographic distribution of conflict for Rwanda and Burundi is mapped in figure 3a. Both Burundi and Rwanda have a long history of ethnic violence. The data on violence comes from Raleigh et al. (2010), and in particular we take the period 1997 (the start of the geocoded data) to 2013 (the time of the experiment). For both Rwanda and Burundi our measure of violence is the sum of conflict-based fatalities in this period within a one decimal degree radius (about 100km) of each respondent's village.

The patterns of blame in regions with different levels of ethnic violence are consistent with previous results (figure 3b). People from high violence regions tend to blame inter-ethnic partners for low offers at a much higher rate than people from low-violence regions. This finding contrasts somewhat with a body of work that suggests that civil conflict *increases* pro-sociality. While the vast majority of reconciliation efforts are aimed towards rebuilding trust and cooperation, our evidence suggests that this misses a considerable part of the story. In east Africa, at least, inter-ethnic trust remains quite high and has been shown to be relatively uncorrelated with violence. Our evidence, instead, suggests that the tendency to blame is heavily correlated with violence, and *much* more common in inter-ethnic partnerships.

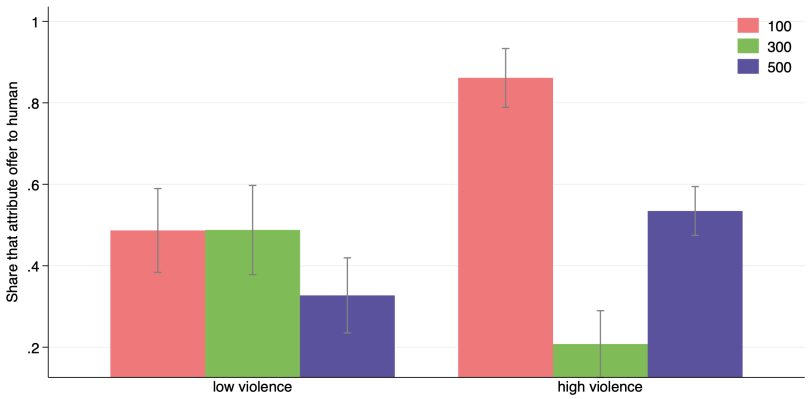
4. Ancillary Hypotheses

We have seen so far that in post ethnic conflict regions, people blame their inter-ethnic partners for surprising outcomes to a far greater extent than is warranted. This is especially pronounced among people for whom ethnicity is particularly salient, and in people still living in the shadow of ethnic violence. Wrongful blame inherently involves confusing noise for bias, however, do we then see that when noise is reduced, that bias also goes down?[12] This hypothesis is related to an argument advanced in a large literature on the contact hypothesis (the seminal contribution is Allport (1954)). According to the contact hypothesis,

[12]Recall that we use bias to refer both to prejudice and to systematic, correlated errors, in the statistical sense. In our case these are the same because the original offers sent by the sender in the victim game are remarkably similar between inter-ethnic and co-ethnic partnerships (see figure 1a). Any differences in ascription on the basis of ethnicity therefore reflects both systematic errors as well as prejudice.



(a) ACELD fatalities 1997-2013



(b) Violence and the Victimhood Bias

Figure 3: The Victimhood Bias and Violence

Note: Figure (a) presents the geographic distribution of violence in the ACLED database, details are in the text. In figure (b) we plot the rate at which individuals hold their partners accountable for the offer received. In both figures we plot the predicted values after regressing the y-axis variable on the controls and the interaction between violence and the offer observed. Violence is divided into high / low based on the median level of violence in the data. Controls include gender, ethnicity, age, age-squared, beliefs, IQ. The sample includes all subjects who were assigned to a partner not from their own ethnic group, and who received a random offer.

when people interact with out-groups more, they get a more precise signal about their behaviour, and accordingly are more likely to view them favourably.

In our case, we see something similar. Table B5 presents results comparing people born in more ethnically diverse villages to those born in relatively homogenous villages. We hypothesize that those with more experience interacting with members of the other community should have less noisy beliefs about inter-ethnic attitudes, and accordingly, should be less likely to place wrongful blame on their partners when something goes wrong. As expected, this is precisely what we find. Those born in an ethnically heterogeneous village are nearly equally likely to assign accountability to their inter-ethnic partner for good and bad outcomes, exhibiting a distribution of ascription similar to both the co-ethnic partnerships, and the true distribution of offers.

Consistent with the vicious cycle inherent in the contact hypothesis, we also see that those who are more likely to blame their partner for bad outcomes, are also the ones *least* likely to select into inter-ethnic partnerships in the first place. We asked respondents to choose five partners that they could be partnered with in a cooperative task that required them to spend time with this person.[13] We used this as a crude measure of willingness to socially interact with those of a different ethnicity.[14]

We find that individuals who chose only those from their own ethnicity displayed much more asymmetry in their ascription than those who chose more diverse sets of potential partners. Indeed in table B6 we see that those who chose the most diverse set of potential partners were among the least likely to blame their inter-ethnic partners for bad outcomes.

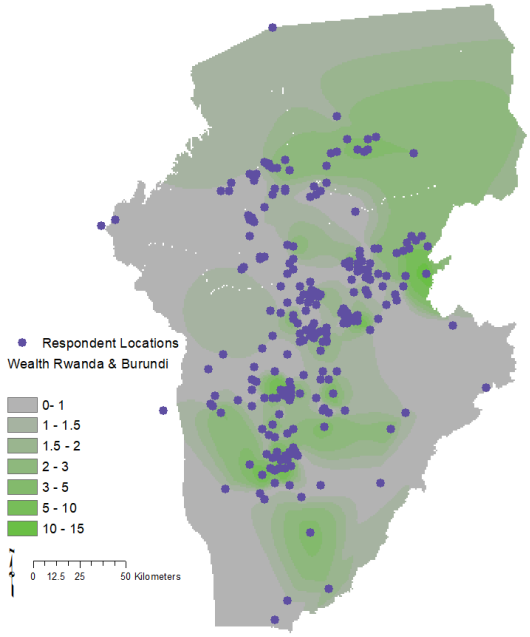
In combination, these two findings highlight a clear rational for policy intervention. Segregation appears quite harmful to the reconciliation process, and those most likely to segregate are the ones we might most wish to target with reconciliation policies.

However, inter-ethnic interaction is not the only mitigating factor. Finally, on the theme of recovery, we examine the role of economic development. We looked at spatial variation in wealth, as proxied by land values.[15] The geographic

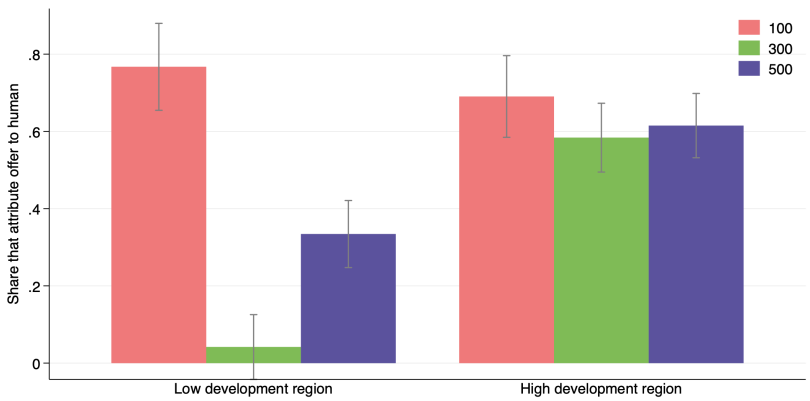
[13] A subset were partnered with someone from their list for the victim game, so the choices were incentive compatible (this is another reason why we focus on the randomized offers only). Also, they chose from those at the session with the caveat that they could not choose anyone they knew or was from their village.

[14] More precisely, we used this to construct a measure of the percentage of their partner choices from the other ethnic group: $Other = \frac{\sum OtherChoices_i}{min\{5, TotalOther\}}$.

[15] Land values are measured with a survey, and thus are self-reported values.



(a) Distribution of Respondent Wealth



(b) The role of development

Figure 4: Development facilitates reconciliation

Note: Figure (a) presents the geographic distribution of wealth in our data, proxied by self-reported lang values. In figure (b) we plot the rate at which individuals hold their partners accountable for the offer received. In both figures we plot the predicted values after regressing the y-axis variable on the controls and the interaction between violence and the offer observed. Wealth is divided into high / low based on the median value of land in the data. Controls include gender, ethnicity, age, age-squared, beliefs, IQ. The sample includes all subjects who were assigned to a partner not from their own ethnic group, and who received a random offer.

distribution of local economic development that we exploit can be seen in figure 4a. Notably, the variation is not simply capturing the huge pocket of (relative) wealth in Kigali - in fact none of our sample is from Kigali, so it appears as grey in the map.

As before, we see some hope for recovery. Economic development brings with it increased economic activity across localized groups, and consistent with this we see a reduced tendency to attribute low offers to partners rather than the dice, and also a much greater likelihood of attributing high offers to partners. In the regions where we might expect people to have the most inter-ethnic interactions, we do see much less inter-ethnic blame for random negative outcomes (figure 4b).

5. Discussion

The difficulty in resurrecting inter-ethnic cooperation in the aftermath of violence and genocide is one of the biggest challenges facing post-conflict societies. This paper uses experimental data collected from the field in Burundi and Rwanda to identify a new behavioural barrier that inefficiently skews the processing of information in a way that constrains inter-ethnic cooperation. Our results show that individuals in inter-ethnic relationships mistakenly attribute negative shocks to deliberate human intent. This mistaking of noise for bias is higher amongst those individuals for whom ethnic identity is salient and those who have been previously exposed to inter-ethnic violence. More broadly, our results suggest that a reduction in noise can end up having a disproportionate impact on inter-ethnic cooperation. This is because the analytical distinction between noise and bias is blurred in how individuals processes information Kahneman et al. (2021).

References

Allport, G. (1954). *The nature of prejudice*. Addison-wesley Reading, MA.

Annan, J., Blattman, C., Mazurana, D., and Carlson, K. (2011). Civil war, reintegration, and gender in northern uganda. *Journal of conflict resolution*, 55(6).

Armand, A., Atwell, P., and Gomes, J. F. (2020). The reach of radio: Ending civil conflict through rebel demobilization. *American Economic Review*, 110(5).

Arrow, K. (1970). Political and economic evaluation of social effects and externalities. In *The analysis of public output*, pages 1–30. NBER.

Arrow, K. J. (1972). Problems of resource allocation in united states medical care. In *The Challenge of Life*, pages 392–410. Springer.

Bauer, M., Fiala, N., and Levely, I. (2014). Trusting former rebels: An experimental approach to understanding reintegration after civil war.

Bellows, J. and Miguel, E. (2009). War and local collective action in sierra leone. *Journal of public Economics*, 93(11-12):1144–1157.

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

Blouin, A. (2022). Culture and contracts: The historical legacy of forced labour. *The Economic Journal*, 132(641):895–951.

Blouin, A. and Mukand, S. W. (2019). Erasing ethnicity? propaganda, nation building, and identity in rwanda. *Journal of Political Economy*, 127(3):1008–1062.

Bohnet, I. and Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 55(4):467–484.

Bratic, V. (2013). Twenty years of peacebuilding media in conflict strategic framework. *UPEACE Open Knowledge Network*, 2:1–36.

Cassar, A., Grosjean, P., and Whitt, S. (2013). Legacies of violence: trust and market development. *Journal of Economic Growth*, 18(3):285–318.

Charness, G. and Levine, D. I. (2007). Intention and stochastic outcomes: An experimental study. *The Economic Journal*, 117(522):1051–1072.

DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., and Zhuravskaya, E. (2014). Cross-border media and nationalism: Evidence from serbian radio in croatia. *American Economic Journal: Applied Economics*, 6(3).

Easterly, W. and Levine, R. (1997). Africa's growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250.

Falk, A., Fehr, E., and Fischbacher, U. (2008). Testing theories of fairness—intentions matter. *Games and Economic Behavior*, 62(1):287–303.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994.

Fershtman, C. and Gneezy, U. (2001). Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics*, 116(1):351–377.

Gächter, S., Kölle, F., and Quercia, S. (2017). Reciprocity and the tragedies of maintaining and providing the commons. *Nature Human Behaviour*, 1(9).

Garcia-Montalvo, J. and Reynal-Querol, M. (2004). Ethnic polarization, potential conflict, and civil wars. *Potential Conflict, and Civil Wars (July 2004)*.

Gilbert, D. T. (1998). Ordinary personology. *The handbook of social psychology*, 2:89–150.

Gino, F., Moore, D. A., Bazerman, M. H., et al. (2008). *No harm, no foul: The outcome bias in ethical judgments*. Citeseer.

Gurdal, M. Y., Miller, J. B., and Rustichini, A. (2013). Why blame? *Journal of Political Economy*, 121(6):1205–1247.

Holmström, B. (1979). Moral hazard and observability. *The Bell journal of economics*, pages 74–91.

Kahneman, D., Sibony, O., and Sunstein, C. R. (2021). *Noise: a flaw in human judgment*. Little, Brown.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., and Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435(7042):673–676.

Kurzban, R., Tooby, J., and Cosmides, L. (2001). Can race be erased? coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, 98(26):15387–15392.

Luis, J., Rowold, D., Regueiro, M., Caeiro, B., Cinnioğlu, C., Roseman, C., Underhill, P., Cavalli-Sforza, L. L., and Herrera, R. (2004). The levant versus the horn of africa: evidence for bidirectional corridors of human migrations. *The American Journal of Human Genetics*, 74(3):532–544.

Mazzocco, P. J., Alicke, M. D., and Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology*, 26(2-3):131–146.

Michalopoulos, S. and Papaioannou, E. (2016). The long run effects of the scramble for africa. *The American Economic Review*, 106(7).

Paluck, E., Porat, R., Clark, C., and Green, D. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, 72(1).

Paluck, E. L. (2012). Media as an instrument for reconstructing communities following conflict. *Restoring civil societies*, pages 284–298.

Paluck, E. L. and Green, D. P. (2009). Deference, dissent, and dispute resolution: An experimental intervention using mass media to change norms and behavior in rwanda. *American Political Science Review*, 103(4):622.

Pettigrew, T. F. (1979). The ultimate attribution error: Extending allport's cognitive analysis of prejudice. *Personality and social psychology bulletin*, 5(4):461–476.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing acled: an armed conflict location and event dataset: special data feature. *Journal of peace research*, 47(5):651–660.

Rao, S. (2014). Sending messages to reduce violent conflict. Technical report, Birmingham, UK: GSDRC.

Rohner, D., Thoenig, M., and Zilibotti, F. (2013). Seeds of distrust: Conflict in uganda. *Journal of Economic Growth*, 18(3):217–252.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In *Advances in experimental social psychology*, volume 10.

Shepard, E. M. and Herrera, R. J. (2006). Genetic encapsulation among near eastern populations. *Journal of Human Genetics*, 51(5):467–476.

Shnabel, N. and Nadler, A. (2008). A needs-based model of reconciliation: Satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. *Journal of personality and social psychology*, 94(1):116.

Shnabel, N. and Nadler, A. (2015). The role of agency and morality in reconciliation processes: The perspective of the needs-based model. *Current Directions in Psychological Science*, 24(6):477–483.

Stangor, C., Lynch, L., Duan, C., and Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, 62(2):207.

Taylor, S. E. and Fiske, S. T. (1978). Salience, attention, and attribution: Top of the head phenomena. *Advances in Experimental Social Psychology*, 11:249–288.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., and Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of personality and social psychology*, 36(7):778.

Appendix A. Data & Protocol Appendix (ONLINE APPENDIX)

Note that this paper uses data that has also been used in Blouin and Mukand (2019) as well as Blouin (2022). This section reviews the protocol and instructions used in those papers, however much of the material in this section already appears in those two papers.

Enumerators were hired from a local country-specific pool used by the firm that we hired to help us manage the data collection. The two data collections efforts, one in Rwanda and another in Burundi, were nearly identical. The main difference was that subjects were directly asked about their ethnicity in Burundi, while this was not permitted in Rwanda. In addition, all instructions (written or oral) were in Kinyarwanda in Rwanda and in Kirundi in Burundi. Data was collected for three projects in mind. The first is Blouin and Mukand (2019) the second is the project described in this paper and the third is a yet unwritten project.

On a given day there was a morning and an afternoon data collection session. Typically the same villages were used for the morning and afternoon sessions. In any given session we typically have 4-5 people from any given village and 4-5 villages present (20 people total), but overall, in the data we have 8-10 subjects from each village (4-5 from the morning and 4-5 from the afternoon).

Data collection sessions took place in a town hall. There was a survey portion and an experimental portion to the session. The surveys took place first, and the experiments took place second. Surveys were completed sitting down at a table in private with a subject. In the experiment portion of the session there 4 were experimental stations and a waiting area. Subjects were in one of these two locations throughout the experiments. In the waiting area there was a large poster board that listed the partnerships for the trust game. The poster board was updated with the offers of the trust game throughout the day if the trust game was assigned to the public treatment.

In each data collection session, there were well-defined roles for our 8 enumerators. They were specialists in that, for example, the person who was in the Enumerator 1 role was in that role for every session. It will help to label these roles as follows, and we'll refer to them as E1-E4:

- Enumerator 1 (1 person)
    1. responsible for greeting subjects as they arrived and handling consent.

2. responsible for matching trust game partners as well as roles (sender / receiver) for the trust game
3. responsible for collecting partner preferences and briefing subjects on how the experiments would run.
4. responsible for managing the flow of experiments, ensuring subjects knew where to go, etc.
5. responsible for payment at the end of the day

- Enumerator 2 (3 people)
  1. responsible for completing surveys with subjects.
  2. responsible for assisting E1 with task 4 above (i.e. logistics / flow of experiments).
- Enumerator 3 (3 people)
  1. responsible for completing surveys with subjects
  2. responsible for completing the trust game with subjects
  3. responsible for completing a separate task (for another paper) with subjects
- Enumerator 4 (1 person)
  1. responsible for completing surveys with subjects
  2. responsible for completing the SIT with subjects (see Blouin and Mukand (2019) for details)

The timeline of events for data collection was as follows:

- Between 8:00 and 8:30am the enumeration team travelled from the hotel in 4 SUVs to the town hall where the data collection session took place.
- The team unloaded materials and started arranging tables, chairs, posters, and other materials needed for the survey and experiments.
- While set-up was taking place each driver drove an SUV to met subjects near their own villages at a pre-determined meeting location, and drove them to the town hall.

- As subjects arrived (typically 4-5 at a time) they were greeted by E1 who described in general terms the survey, experiments and the purpose of the study before distributing and reading the consent agreement.

- After collecting the consent agreements, each subject was given an ID card that they pinned to their shirt which listed a letter that corresponded to the region they live in (a geographic cluster close to the pick-up location), and a number identifying each individual from that region. ID tags were in bags corresponding to the region (i.e. letter) and were dolled out randomly upon subject arrival (conditional on letter).

- As they entered the town hall, each subject was paired up with any one of seven enumerators (i.e. E2-E4) to complete the survey. These enumerators simply lined up near the entrance of the town hall and paired-up with subjects as they entered.

- As subjects were completing the surveys with one of the other enumerators, E1 matched each subject to another for the trust game and determined roles (who was to play as sender / receiver). This matching was done using numbers from the ID tags, and was done without any consideration (or knowledge of) the ethnic identity of the subjects. At this stage the only consideration when matching two subjects was logistical - to ensure that subjects who were matched to each other were not from the same village. Once E1 knew the exact composition of the session (i.e. how many subjects from each village) she ensured that, for instance, whoever had ID tag A2 would be partnered with whoever got B5. An 'A' (i.e. from region A) was never partnered with another 'A', and likewise for every other letter. So as a first step, we assumed that if two people were not from the same village it was unlikely they would know each other. This matching process typically took no more than 15 minutes.

  - The only reason this was not completely trivial (and hence why it typically took more than 30 seconds) is each subject could only be partnered with another individual once, to prevent outcomes from one game influencing another. For example, subjects played the trust game twice, once as a sender and once as a receiver. E1 always made sure that that any partnership only ever took place once.

- As subjects finished the surveys, they were sent back to E1. She briefed

each subject (sequentially and one at a time) on logistics and also elicited information in the following order:

- Partner-selection task: at this point the partner-preferences were taken for each subject. Each subject was asked to look around at the ID tags of the people at the session, and list the top 5 individuals that they would like to be partnered with to take part in a cooperative task. They were told that a few people would be matched with a partner of their choice for the last game of the day (which was the task under E3 item 3 in the list of responsibilities above)
- Payment Protocol. Each subject was informed that the monopoly money he/she received, represented real money. Every dollar of monopoly money represented a Rwandan or Burundian Franc. At the end of each experiment an enumerator would write down on a piece of paper how much money was earned in the experiment, signed the back of the piece of paper, and put the piece of paper in a sac that the subjects were responsible for. At the end of all experiments, subjects would reach into the sac and pull out one piece of paper, and would be paid in cash, the amount listed. It was stressed that total money earned would depend on outcomes of the various lab exercises, but that they would earn money from only one specific exercise, chosen at random.
- Question on previous acquaintance: As noted above, subjects were assigned trust-game partners randomly. Since the partner pairings were done while surveys were being completed, at this point subjects were informed of who their partner would be. The enumerator asked subjects (individually) if they had ever met their partner before. If either had, the protocol dictated that new partners be found.

- Subjects were then taken to an area of the town hall that had seats for them to wait until they were called to participate in an experiment. They often had to briefly wait for an assigned partner to complete the survey.
- As subjects became available, E1 called out the ID tags of subjects and brought each subject them over to E3 or E4, at one of the experiment stations depending on who was free.
- When a subject arrived at E4's work station, he implemented the SIT for the subject (see Appendix F).

- In the case of the trust-game, as subjects arrived and sat down at the trust-game station, E1 (or E2) told E3 which subject was player 1 and which was player 2.[16] This is to to ensure that each player got to play once as 'sender' and the other time as 'receiver'.

  – E3 explained the experiment to both subjects and gave out the Monopoly money to player 1. There was a script that they read, and then they asked both subjects if they understood both the rules and the implications of the decisions. The enumerator was free to explain the trust game in their own words, if they felt that subjects did not understand the game after being read the script.

- After playing the trust game, the subject was sent back to the waiting area until they could participate in another experimental exercise.

- Part of the job of E1 and the three E2s was to keep an eye out for idle experimental stations and subjects that had not yet participated in those stations. E.g. if E4 was idle and a participant in the waiting area had not completed the SIT, E1 or E2 would be responsible for making sure that E4 completed the SIT with the subject at that time.

- Once a subject finished all the experiments, they were called from the waiting area, and they pulled from their sac a piece of paper that listed their payment. They were given this payment plus a participation fee plus a case of soap. They returned the ID tag and were free to leave.

- As they exited the town hall, the drivers were waiting for them to take them back to their village. When drivers dropped off one set of subjects another set was waiting at the same location for the afternoon session.

- While drivers dropped off and picked up subjects, enumerators typically went into town to have lunch. In the evening session when drivers dropped off subjects, the enumerators organized all of the surveys and experimental materials. Drivers returned, the SUV was packed up, and the team typically went for dinner near the hotel.

[16] Recollect that after the survey was completed, all the E2's had only one responsibility - namely to help E1 with logistics.

A.1. Original Experiment Instructions:

Note that this paper uses data that has also been used in Blouin and Mukand (2019) as well as Blouin (2022). This section reviews the protocol and instructions used in those papers, however much of the material in this section already appears in those two papers.

Q1) Player 1 ID:_____

Q2) Player 2 ID:_____

For all respondents, read the following:

Here's how the exercise works: Player 1 receives 600RWF. You can put some money onto the table. I'm going to roll a dice. If the dice shows a 4, 5, or 6, player 2 will receive the money put into the envelope by player 1. If the dice shows a 1, 2, or 3 the offer will be randomly determined. Whatever is placed on the table will be doubled. Then Player 2 gets to decide how divide all of the money on the table between the two of you.

Enumerator instructions:

- Make sure they understand the game.
- Explain in your own words if necessary.
- Give the 6 bills of Monopoly money to player 1.

Q3) Player 1: How much of your 600RWF would you like to share? *Circle one*

a) 0 RWF
b) 100 RWF
c) 200RWF
d) 300RWF
e) 400RWF
f) 500RWF
g) 600RWF

Instruction to Enumerator: Now roll a dice ***privately*** *so that neither player can see. If the roll is a 1 give player 2 200RWF. If the roll is a 2 give player 2*

600RWF. If the roll is a 3 give player 2 1,000RWF. If the offer is a 4, 5, 6 double what was given by player 1 and give the money to player 2.

Player 2: I have taken the money shared with you and I have doubled it. You can now decide how to divide this money between the two of you. How much would you like to keep and how much would you like player 1 to have, in addition to what (s)he has already kept?

Q4) Player 2 keeps:_____

Q5) Player 2 shares with player 1: _____

Q6) Player 2: These tokens represent the amount of money that has been shared with you. Do you think this amount was randomly chosen or do you think it was chosen by player 1? If you guess right, you will earn an additional 100RWF.

Circle One

a) Randomly Chosen

b) Player 1 Chose

Enumerator Instructions: Now write down on a 'payment stub' for each person, how much that person earned in the experiment. Sign or initial the back of the stub and place it in the sac that they have to hold the stubs.

Appendix B. Additional Evidence (ONLINE APPENDIX)

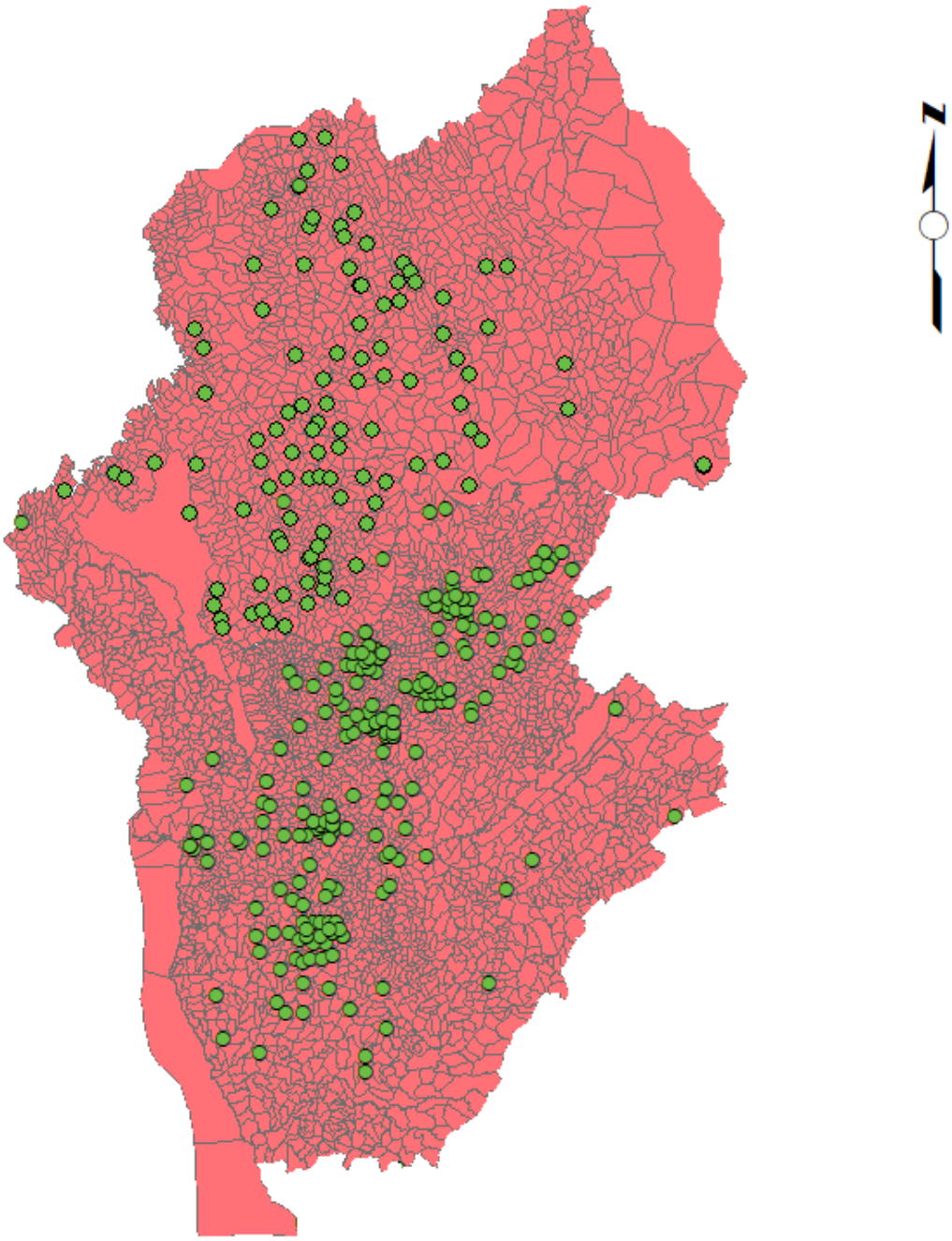


Figure B1: Respondent Locations

Table B1: Summary Statistics and Balance

	Means			Medians		
	Panel A: Balance by partner assignment					
	Co-ethnic	Inter-ethnic	p	Co-ethnic	Inter-ethnic	p
Attribution to Human	.597	.563	.341	1	1	.341
Tutsi	.136	.495	6.32e-30	0	0	0
Gender	1.38	1.43	.178	1	1	.178
Age	41.2	42	.372	40	41	.446
Income	213,419	266,797	.15	100,000	100,000	.98
Education	5.72	5.29	.0829	6	6	.0396
Rational belief	.827	.848	4.88e-09	.857	.857	6.78e-10
Raven score	.645	.64	.793	.75	.75	.998
	Panel B: Balance by low offer					
	Low Offer	Not low offer	p	Low Offer	Not low offer	p
Attribution to Human	.576	.608	.457	1	1	.457
Tutsi	.287	.268	.624	0	0	.624
Gender	1.4	1.39	.685	1	1	.684
Age	42.2	39	.00502	41	39	.00575
Income	229,311	256,198	.543	100,000	100,000	.211
Education	5.53	5.6	.835	6	6	.97
Rational Belief	.836	.833	.561	.857	.857	.65
Raven Score	.636	.67	.144	.75	.75	.33

Note: The table plots descriptive statistics. We highlight balance across inter-ethnic and co-ethnic samples in panel A and balance across low / high offers in panel B. Note that in panel A Tutsi is mechanically unbalanced since Tutsi make up a smaller share of each country, and beliefs regarding ascription are based on offers, and are endogenously unbalanced. Otherwise we see broad balance of covariates, consistent with random assignment. Only education differs between co-ethnic and inter-ethnic samples in panel A, which is a function of inter-ethnic differences in levels of education. Only age differs between by offer in panel B. Of the 13 attributes x treatment cells where we expect balance, we see statistically significant differences in 1, which is broadly consistent with random assignment.

Table B2: Regression estimates: inter-ethnic pairings

	(1)	(2)	(3)	(4)
	Likelihood of holding partner accountable for outcome			
Low offer	0.0311	0.0306	0.0332	0.0339
	(0.0647)	(0.0652)	(0.0675)	(0.0678)
Inter-ethnic game	-0.128*	-0.128*	-0.145*	-0.142*
	(0.0672)	(0.0678)	(0.0729)	(0.0754)
Low offer x Inter-ethnic game	0.287***	0.288***	0.290***	0.288***
	(0.0799)	(0.0806)	(0.0843)	(0.0852)
Individual baseline characteristics	✓	✓	✓	✓
IQ score (Raven)		✓	✓	✓
Rational belief (see eqn. 2)			✓	✓
Education				✓
Observations	330	330	330	330
R-squared	0.152	0.153	0.159	0.159
Dependent Variable Mean	0.542	0.542	0.542	0.542

Note: The table regresses the rate at which individuals hold their partners accountable for the offer received on a set of controls and the interaction between receiving a low offer and being assigned to an inter-ethnic partnership. We include individual baseline characteristics gender, age, age-squared, ethnicity, and income. In subsequent columns we add IQ as measured by a Raven test, their rational beliefs about whether the offers came from a person (described in detail in the text), and their education level. Standard errors are clustered at the lab-session level. *** denotes statistical significance at the 1% level; ** denotes statistical significance at the 5% level; * denotes statistical significance at the 10% level. The sample includes all respondents that received a random offer in the victim game.

Table B3: Regression estimates: ethnic salience

	(1)	(2)	(3)	(4)
	Likelihood of holding partner accountable for outcome			
Low offer	0.177	0.175	0.176	0.167
	(0.124)	(0.123)	(0.125)	(0.126)
High ethnic salience	-0.0662	-0.0894	-0.0869	-0.0964
	(0.125)	(0.125)	(0.128)	(0.126)
High ethnic salience x Low offer	0.363**	0.378**	0.379**	0.390**
	(0.167)	(0.163)	(0.163)	(0.165)
Individual baseline characteristics	✓	✓	✓	✓
IQ score (Raven)		✓	✓	✓
Rational belief (see eqn. 2)			✓	✓
Education				✓
Observations	139	139	139	139
R-squared	0.262	0.265	0.266	0.273
Dependent Variable Mean	0.525	0.525	0.525	0.525

Note: The table regresses the rate at which individuals hold their partners accountable for the offer received on a set of controls and the interaction between receiving a low offer and being ethnic salience. Ethnic salience is divided into high / low based on the median value of salience in the sample. We include individual baseline characteristics gender, age, age-squared, ethnicity, and income. In subsequent columns we add IQ as measured by a Raven test, their rational beliefs about whether the offers came from a person (described in detail in the text), and their education level. Standard errors are clustered at the lab-session level. *** denotes statistical significance at the 1% level; ** denotes statistical significance at the 5% level; * denotes statistical significance at the 10% level. The sample includes all respondents that received a random offer in the victim game and were assigned to an inter-ethnic partner in the victim game.

Table B4: Regression estimates: ethnic violence

	(1)	(2)	(3)	(4)
	Likelihood of holding partner accountable for outcome			
Low offer	0.188	0.181	0.182	0.175
	(0.122)	(0.117)	(0.117)	(0.118)
High violence	0.0261	0.0245	0.0251	0.0281
	(0.0971)	(0.0995)	(0.101)	(0.102)
Low offer x high violence	0.235	0.263*	0.263*	0.267*
	(0.153)	(0.144)	(0.143)	(0.146)
Individual baseline characteristics	✓	✓	✓	✓
IQ score (Raven)		✓	✓	✓
Rational belief (see eqn. 2)			✓	✓
Education				✓
Observations	139	139	139	139
R-squared	0.254	0.260	0.260	0.268
Dependent Variable Mean	0.525	0.525	0.525	0.525

Note: The table regresses the rate at which individuals hold their partners accountable for the offer received on a set of controls and the interaction between receiving a low offer and living in a village exposed to high rates of violence. Violence is divided into high / low based on the median value of violence in the sample. We include individual baseline characteristics gender, age, age-squared, ethnicity, and income. In subsequent columns we add IQ as measured by a Raven test, their rational beliefs about whether the offers came from a person (described in detail in the text), and their education level. Standard errors are clustered at the lab-session level. *** denotes statistical significance at the 1% level; ** denotes statistical significance at the 5% level; * denotes statistical significance at the 10% level. The sample includes all respondents that received a random offer in the victim game and were assigned to an inter-ethnic partner in the victim game.

Table B5: Regression estimates: ethnic diversity

	(1)	(2)	(3)	(4)
	Likelihood of holding partner accountable for outcome			
Low Offer	0.408**	0.395**	0.395**	0.410**
	(0.148)	(0.149)	(0.150)	(0.165)
Heterogenous region	-0.117	-0.142	-0.142	-0.155
	(0.114)	(0.120)	(0.123)	(0.130)
Low Offer x Heterogenous region	-0.471**	-0.420*	-0.418*	-0.436*
	(0.215)	(0.212)	(0.209)	(0.215)
Individual baseline characteristics	✓	✓	✓	✓
IQ score (Raven)		✓	✓	✓
Rational belief (see eqn. 2)			✓	✓
Education				✓
Observations	84	84	84	84
R-squared	0.222	0.239	0.239	0.248
Dependent Variable Mean	0.476	0.476	0.476	0.476

Note: The table regresses the rate at which individuals hold their partners accountable for the offer received on a set of controls and the interaction between receiving a low offer and being born in a village with high ethnic diversity. Diversity is divided into high / low based on the median value of diversity in the sample. We include individual baseline characteristics gender, age, age-squared, ethnicity, and income. In subsequent columns we add IQ as measured by a Raven test, their rational beliefs about whether the offers came from a person (described in detail in the text), and their education level. Standard errors are clustered at the lab-session level. *** denotes statistical significance at the 1% level; ** denotes statistical significance at the 5% level; * denotes statistical significance at the 10% level. The sample includes all respondents that received a random offer in the victim game and were assigned to an inter-ethnic partner in the victim game. It also only includes subjects from Rwanda, because we only have ethnicity shares by village in Rwanda.

Table B6: Regression estimates: ethnic partner preferences

	(1)	(2)	(3)	(4)
	Likelihood of holding partner accountable for outcome			
Low offer	0.476***	0.485***	0.485***	0.474***
	(0.104)	(0.105)	(0.105)	(0.105)
Chose ethnic heterog team	0.138	0.139	0.139	0.126
	(0.0857)	(0.0865)	(0.0871)	(0.0881)
Low offer x Chose ethnic heterog team	-0.348**	-0.354**	-0.354**	-0.335**
	(0.162)	(0.159)	(0.160)	(0.162)
Individual baseline characteristics	✓	✓	✓	✓
IQ score (Raven)		✓	✓	✓
Rational belief (see eqn. 2)			✓	✓
Education				✓
Observations	139	139	139	139
R-squared	0.262	0.265	0.265	0.269
Dependent Variable Mean	0.525	0.525	0.525	0.525

Note: The table regresses the rate at which individuals hold their partners accountable for the offer received on a set of controls and the interaction between receiving a low offer and selecting an ethnically heterogeneous set of potential lab partners. Partner preference is divided into high / low based on the median value of diversity-preference in the sample. We include individual baseline characteristics gender, age, age-squared, ethnicity, and income. In subsequent columns we add IQ as measured by a Raven test, their rational beliefs about whether the offers came from a person (described in detail in the text), and their education level. Standard errors are clustered at the lab-session level. *** denotes statistical significance at the 1% level; ** denotes statistical significance at the 5% level; * denotes statistical significance at the 10% level. The sample includes all respondents that received a random offer in the victim game and were assigned to an inter-ethnic partner in the victim game.

Table B7: Regression estimates: economic development

	(1)	(2)	(3)	(4)
	Likelihood of holding partner accountable for outcome			
Low offer	0.429***	0.439***	0.441***	0.445***
	(0.0927)	(0.0916)	(0.0896)	(0.0908)
High development region	0.208**	0.218**	0.224**	0.215**
	(0.0903)	(0.0877)	(0.0918)	(0.0944)
Low offer x High development region	-0.257	-0.263*	-0.274*	-0.292*
	(0.155)	(0.153)	(0.152)	(0.157)
Individual baseline characteristics	✓	✓	✓	✓
IQ score (Raven)		✓	✓	✓
Rational belief (see eqn. 2)			✓	✓
Education				✓
Observations	139	139	139	139
R-squared	0.260	0.265	0.265	0.271
Dependent Variable Mean	0.525	0.525	0.525	0.525

Note: The table regresses the rate at which individuals hold their partners accountable for the offer received on a set of controls and the interaction between receiving a low offer and being from a high-welath village. We proxy for wealth using self-reported land values, and wealth is divided into high / low based on the median land value in the sample. We include individual baseline characteristics gender, age, age-squared, ethnicity. We omit income from our baseline controls because of the high co-linearity with land-values. In subsequent columns we add IQ as measured by a Raven test, their rational beliefs about whether the offers came from a person (described in detail in the text), and their education level. Standard errors are clustered at the lab-session level. *** denotes statistical significance at the 1% level; ** denotes statistical significance at the 5% level; * denotes statistical significance at the 10% level. The sample includes all respondents that received a random offer in the victim game and were assigned to an inter-ethnic partner in the victim game.