Patrick Nyheim Schjølberg

Distrusting Ethnic Changes?

A quantitative study on the moderating role of ethnic group change in the refugee population for refugee-hosting communities' social trust in Uganda, Tanzania, and, Zambia

Masteroppgave i statsvitenskap
Veileder: Karin Dyrstad
Medveileder: Siri Aas Rustad
Juni 2023

Masteroppgave

NTNU
Norges teknisk-naturvitenskapelige universitet

NTNU
Kunnskap for en bedre verden

Patrick Nyheim Schjølberg

Distrusting Ethnic Changes?

A quantitative study on the moderating role of ethnic group change in the refugee population for refugee-hosting communities' social trust in Uganda, Tanzania, and, Zambia

Masteroppgave i statsvitenskap
Veileder: Karin Dyrstad
Medveileder: Siri Aas Rustad
Juni 2023

Norges teknisk-naturvitenskapelige universitet

NTNU
Kunnskap for en bedre verden

Abstract

The effects refugees have on African host-community attitudes toward social trust are poorly studied. The mechanisms of this relationship are largely derived from studies on European states (Dinesen & Sønderskov, 2015; Jeannet, 2020; McLaren, 2017). This is troubling, as the relative number of refugees far exceeds that of European countries. Few studies have tackled the relationship between trust and refugees quantitatively with African survey data, and those who have are usually concerned with political and institutional trust (e.g. Hutchison & Johnson, 2011). In this thesis, *generalized social trust* is studied, which is often revered as fundamental for larger society to function well, and in some traditions a key determiner of political trust. As African countries display far lower levels of social trust than European countries, the mechanisms of how trust both develops and is maintained may be different from the high-trust societies of the Global North.

Ethnicity is likewise an area in which African countries greatly differ from the ethnically quite homogeneous European nation-states. Ethnicity is viewed as an important determiner of how social trust develops and manifests. The ethnicity of refugees, ethnic fractionalisation in the host community, and ethnic differences between refugees and host-community members can therefore be similarly important yet stands as previously unexplored intermediary explanations for understanding how refugees affect social trust. This thesis makes an important first step by utilising a unique difference-in-difference-in-differences design in Uganda, Tanzania, and Zambia framed by the developments in the Kivu conflict in the Democratic Republic of the Congo. This thesis analyse how proximity to refugees affect trust if the majority ethnic group changes compared to where it is constant. The results indicate that living in proximity to a refugee settlement may be associated with higher social trust, whereas proximity to settlements where ethnic change occurs in the refugee group have a strong and statistically robust negative effect on social trust for the host-community members.

Sammendrag

Effekten flyktninger har på sosial tillit i afrikanske vertssamfunn er understudert. Mekanismene for forholdet beror hovedsakelig på studier av europeiske land (Dinesen & Sønderskov, 2015; Jeannet, 2020; McLaren, 2017). Dette er problematisk, da flyktningsituasjonen i Afrika sør for Sahara er langt mer omfattende. Få studier har kvantitativt undersøkt forholdet mellom tillit og flyktninger med afrikansk data, og det lille som eksisterer har hovedsaklig fokusert på politisk og institusjonell tillit (f.eks. Hutchison & Johnson, 2011). I denne avhandlingen undersøker jeg *generalisert sosial tillit,* som anses som fundamentalt for et velfungerende samfunn. Siden afrikanske land har langt mindre sosial tillit enn europeiske land, kan mekanismene for hvordan tillit utvikles og opprettholdes være forskjellige fra høytillitssamfunnene i den globale norden. Etnisitet er likeledes et område der afrikanske land i stor grad skiller seg fra de etnisk mer homogene europeiske nasjonalstatene.

Etnisitet anses som en viktig faktor for hvordan sosial tillit utvikler og manifesterer seg. Etnisiteten til flyktninger, etnisk fragmentering i vertssamfunnet og etniske forskjeller mellom flyktninger og medlemmer av vertssamfunnet kan derfor være avgjørende, men står samtidig som uutforskede forklaringer på hvordan flyktninger påvirker sosial tillit. Denne avhandlingen tar et viktig første skritt ved å benytte et unikt difference-in-difference-in-differences design med casene Uganda, Tanzania og Zambia. Dette med basis i utviklingen i Kivu-konflikten i Den demokratiske republikken Kongo. Avhandlingen analyserer hvordan nærhet til flyktninger påvirker tillit når flyktninggruppen blir utbyttet av en gruppe med annen etnisitet, sammenlignet med tilfeller der en slik endring ikke finner sted. Resultatene antyder at å bo i nærheten av flyktningleirer kan være assosiert med høyere sosial tillit, mens nærhet til leirer der en etnisk utskiftning i flyktninggruppen finner sted har en sterk og statistisk signifikant negativ effekt på sosial tillit hos respondentene i vertssamfunnet.

Acknowledgements

First and foremost, I want to thank my partner, Nina, for always believing in me, especially when I don't. I wouldn't be who I am today if it wasn't for you, and neither would this thesis—and that's not only due to the numerous readthroughs and discussions. You inspire me and those around you, I couldn't be happier with a partner. Thank you.

I want to thank my small but no less important family. I especially want to thank my mother for relentless support throughout the years and always cheering me on when I most need it, I always cherish words, though it may not always seem so.

I want to thank Halvard Buhaug and Karin Dyrstad for giving me the opportunity to write this MA-thesis as part of the TRUST project, and I want to thank Siri Rustad for seeing me through it. You and Karin have provided much necessary steering of an otherwise stupidly ambitious project, and I am truly happy that I have taken your great insights to heart. I also want to thank you for the numerous opportunities you have provided, both past and future. You are an important reason for me wanting to continue with research going forward.

I want to thank the members of the VIP group at NTNU for having me both as a participant and for the fantastic advice I received at my presentation. Being part of a research group has been a wonderful experience and the seminars have given me so much. I want to issue a special thanks to Micheal Kumove and Sofie Hjorthol Grønset for reaching out and providing great discussion and advice, and to Charles Butcher and Thea Johansen for putting their academic faith in me.

I want to thank the members of the TRUST project for great feedback, and for the financial support received whilst working on the thesis. I am looking forward to meeting you in person and continuing research on this important topic together come autumn.

I want to thank Pausegjengen at 9457 for the great breaks, they have arguably been both too great and too long. Never have I met people that so quickly and openly took me in. I especially wish to offer my gratitude for the new friendships I have found in Oda, Jørgen, Emil, and Anna both academically and socially. Y'all better come to Oslo ASAP. There are a lot of people I am grateful for, and if it weren't for them, I wouldn't have wound up precisely with this MA thesis. Thanks for the lifelong friendships, and for the many small but important meets.

-Patrick Schjølberg, 2023

Table of Contents

1. Introduction

Forced migration is at an all-time high since World Word II. As of 2021, over 83 million people were forced to flee their homes. Of these, more than a quarter resides in Sub-Saharan Africa (United Nations High Commissioner for Refugees (UNHCR), 2021). Immigration remains a core political issue in Europe, despite that the relative number of refugees remains much smaller than in Sub-Saharan Africa. This is exemplified by the currently halted UK-Rwanda refugee scheme, where *all* illegal migrants arriving in the UK were supposed to be outsourced to Rwanda on a one-way ticket against monetary compensation for the Rwandan government. This despite Rwanda already hosting twice the number of refugees relative to the UK. Such schemes may present a political solution for Western governments facing anti-immigration pressure and can be economically beneficial for African governments seeking monetary sovereignty (Easterly, 2007). However, it is at the individual and meso-level the effects of such schemes will have adverse effects. Existing forced migration patterns already put a severe amount of stress on a host of Sub-Saharan African countries. European manipulation of the migration patterns can compound the situation.

Mirroring how the western public understands the refugee crises, research on the effects refugees have on host-society is largely European centred. There are glaringly few empirical studies on the effects of refugees on attitudes of trust in the host community in the Global South, despite the refugee situation being far greater. Untangling the mechanism of how refugees impact Sub-Saharan host-communities are pivotal to understanding the effects of the growing refugee-crisis, finding solutions, and adapting policy. This thesis is written as a response to this empirical gap, as part of the PRIO TRUST project and its stated goals. I ask how Sub-Saharan Africa differs from the European context in this regard, and why European insights may not be exportable.

African countries often constitute what Fukuyama coins as “low-trust equilibriums” (2015). Only 14% of Afrobarometer respondents believed that most people could be trusted in 2021 (Logan & Torsu, 2022). Generalised social trust is often viewed as necessary for positive economic, social, and democratic growth. Moreover, Africa’s colonial past with haphazardly drawn borders still haunts the continent to this day, instead of the pre-colonial more relevant borders (Wishman & Butcher, 2022). Ethnic composition of society in Sub-Saharan Africa is severely different from the largely Western European nation-state prototype. Ethnic diversity have shown itself as a central, but debated, component in understanding both provision of public goods, conflict, and economic growth (Ellingsen, 2011; Habyarimana et al., 2007; Posner, 2004; Turton, 1997). A substantial literature has analysed the link between ethnic diversity and social trust; however, the insights are largely drawn from studies on the Global North. The presence and arrival of refugees can influence this ethnic diversity and composition of host society (Fisk, 2019; Rüegger & Bohnet, 2018). The idea that the ethnicity of refugees is similarly important and can have an effect on social trust in the host-community is cogent. This thesis builds from these assumptions and asks:

How does refugees affect generalized social trust in Sub-Saharan African host communities, and is this relationship moderated by changes in the majority ethnic refugee group?

This research question consists of two components and will be analysed stepwise with a quasi-experimental difference-in-differences design. Firstly, it asks how the presence of refugees, regardless of ethnicity, affects generalised social trust—that is, trust in strangers—in the host community. Secondly, I theorise that this relationship will be affected when the ethnic refugee group changes in the host community. By utilising a framing conflict for case selection, precisely such a situation occurs in Uganda, Tanzania, and Zambia.

I utilise a lean and simple theoretical framework that is prevalent in most trust-based empirical studies, and a likewise simple theoretical approach for ethnicity to build the argument. I use this framework in combination with the few similar empirical studies in an African context to outline simple, yet powerful hypotheses.

In contrast to this simplicity, empirically investigating this in an African context requires a fair bit of methodological ingenuity, and the overall focus of the thesis reflects this. I mainly base my analysis on Afrobarometer survey data from the host-communities, but this is insufficient for both parts of the research question, particularly pertaining to ethnicity. I employ a quasi-experimental quantitative difference-in-differences design to test both the effects of refugees on social trust, as well as the viability of ethnic group change moderation on this relationship. This requires a modification of the triple difference estimator, which I develop in the paper. I analyse how the relationship varies between the cases with temporal and spatially operationalised data. This design is not without its caveats: analysing this question in an African context is problematic due to both the availability of data, and the quality of this data. I therefore draw upon six unique datasets to fully address both parts of the research question.

This comes with its own set of problems, particularly for data overlap and ability to combine datasets. To overcome central empirical and methodological issues, I base the thesis around a framing conflict. This limits the cases of study and helps mitigate a series of empirical challenges, such as identifying the ethnicity of refugees and ascertaining time of arrival to be after the first measurement. I base this on the Kivu conflict in the Democratic Republic of Congo (DRC), erupting in 2006. The most natural countries to receive refugees following the conflict are the closest neighbouring countries (Rüegger & Bohnet, 2018). The countries that have participated in Afrobarometer will therefore serve as the cases: Uganda, Tanzania, and Zambia, as they have overlapping data from all the employed datasets. This ensures both before and after measurements of social trust at the individual level and enables construction of multiple treatment variables proxying for exposure to different refugee groups. The three cases combined with the framing conflict puts the thesis between a rock and a hard place in terms of methodology. This thesis builds on the Kivu conflict development and descriptive data on the ethnicity of the DRC refugees to create treatment variables employed in quantitative large N respondent survey data analysis.

The analysis of social trust in the host community is made possible with spatial separation of respondent's administrative districts with and without refugee settlements. The results indicate an empirical puzzle: the presence of refugees overall may be associated with a positive effect on social trust in the host society. However, this effect is heavily and significantly moderated by ethnic change in the refugee group. The change in the majority ethnic group for the DRC refugee settlements is strongly associated with a substantial decline in social trust in the host-community. This is particularly interesting, as in both Uganda and Zambia these are the districts that display some of the highest levels of social trust before the outbreak of the Kivu conflict and subsequent arrivals of refugees from a different ethnic group than prior to the conflict.

The results are robust on a variety of models, alternate variables, and controlled with socio-economic explanatory variables. The results are consistent with qualitatively identified ethnic changes in the cases. Overall, this thesis presents a new way of quasi-experimentally analysing the effects of social trust in the host community in Sub-Saharan Africa and highlights the explanatory and important moderating effect of ethnic changes in the refugee group.

1.2 Thesis Structure

A recurring theme is that the thesis operates in mostly unchartered waters without clear theoretical framework and similar prior research. As such, the thesis manoeuvres in and between a series of fields and draws on them theoretically and empirically for both research design, datasets, and the building of assumptions and hypotheses. The literature on social trust, ethnicity, and migration is particularly central. I outline the core concepts of these fields in chapter 2, which frames the discussion of central—mainly European—literature in chapter 3. As this is an area of limited research in an African context, the few empirical studies with similar research question are heavily utilised in the construction of the theoretical framework, assumptions, and hypotheses of Chapter 4. This chapter theorises how the relationship of social trust and refugees may be different in a Sub-Saharan context, particularly due to ethnic composition. Chapter 5 presents the overarching research design, illuminating the importance of finding suitable cases via the framing conflict. The framing conflict as well as the refugee populations and ethnicity of refugees in the subsequent cases, Uganda, Tanzania, and Zambia, is presented in chapter 7.

Chapter 8 is the most substantial part of the thesis, and is devoted to data, data wrangling, and the many steps taken to operationalise treatment variables for the difference-in-difference design. The research design calls for substantial and complex temporal and spatial data. This chapter is followed by chapter 9 on statistical methods, which presents the core empirical strategy for the stepwise quantitative analysis. Particularly important is here the difference-in-difference design as well as my own modified version of the triple difference estimator necessary for testing the moderating effect of ethnic change in the refugee group. Chapter 10 presents the results of the stepwise analysis, followed by chapter 11 that discusses the implications, shortcomings, and insights of the analysis before chapter 12 concludes the thesis. Supplementary material and additional tables are included in the appendix following the literature chapter.

2. Conceptual framework

Before reviewing the literature on social trust, migration, and ethnicity, the core concepts are best presented first. These sections cover the central theoretical arguments as well as why these elements are central to understanding the effect hosting refugees have for the host-community in Sub-Saharan Africa. The discussion on social trust is treated the most closely here, whereas both ethnicity and migration is foremost expanded on—in relation to social trust—in the theoretical framework chapter succeeding the literature review.

2.1 Social trust

Fundamental to the thesis is the conceptualisation and understanding of trust. In a way, trust is a simple, straightforward concept that is utilised in everyday speech. The capacity to trust one's fellow person is largely *felt* and by extension also intuitively relatable and understandable. In this way, trust is largely something that exists for individuals at varying degrees. Yet, trust is highly contextual and therefore relative. You might trust a colleague to send you documents on time, though you might not trust the same colleague with taking care of your children over the weekend. Trust is dependent on the matter of the trust; your prior experiences with the individual; and your prior experiences with similar situations. Understood this way, trust becomes a far more complex concept—a complexity that has spurred a rich and lively debate in the literature. Delhey & Newton encapsulates this essence: "For all that has been written about it in recent years, there is no general theory of trust. Rather, there is a degree of conceptual confusion, and a variety of partial approaches" (2005, p. 312). Perhaps it was precisely this conceptual confusion that lead Newton to write the widely referenced chapter on trust in *The Oxford Handbook of Political Behaviour* (2007). Newton defines trust as "the belief that others will not deliberately or knowingly do us harm, if they can avoid it, and will look after our interests, if this is possible" (Newton, 2007, p. 333-334). Albeit vague, it captures the two major elements of trust as either positive or negative. Negative, as in trusting that others will not deliberately do us harm, and positive, that others will ensure our interests.

2.2 Types of Trust

The most common differentiation of trust is between *political* and *social* trust. Political trust is often understood as vertical trust, as trust in institutions, political leaders, religious leaders, police. I.e., it is not between individuals, but the individual and most commonly the state (Newton, 2007). This thesis is concerned with horizontal trust, which is between members of society, i.e., the people around you. Trusting that the state—broadly understood—is neither out to harm you, but also maintains your interests, is different than trusting the *people around you* in the same manner. If one views human beings as inherently well-meaning and good, this will likely translate to higher horizontal trust. That same individual might not view the state in a similar manner: the state is more than the sum of its human parts.

In political science, a clear preference for analysing political trust as opposed to social trust is evident. As a concept of study, social trust departs somewhat from *political* science. The seeming neglect of social trust is problematic in several ways. First, how social and political trust affects one another is a contested topic. Fukuyama argues in *Political Order and Political Decay* (2015) that political trust partly stems from the level of social trust in society. In particular, social trust is commonly theorised to be a key determinant of political trust in the so-called cultural explanations of social trust. In this tradition, political trust is seen as an extension of interpersonal trust, building on the cultural norms and beliefs about people (Mishler & Rose, 2001). Zmerli & Newton (2008) find statistical evidence for such a relationship. However, the more recent study by Sønderskov & Dinesen (2016) find that trust in state institutions has a causal impact on social trust and little evidence of the former. The overemphasis on analysing political trust without taking social trust into account may therefore lead to flawed insights. This thesis seeks in part to remedy this neglect by furthering the understanding of how *social trust* in the host-community is affected by refugee presence, specifically in the less studied African context.

Causation aside, social trust differs from political trust in other ways as well. One central topic in the literature is if there exist multiple dimensions of social trust, or if it analytically should be treated as a single entity. Despite some evidence for the latter (Whiteley, 2000), it has become standard to distinguish between two broad types of social trust: *Particularised* and *generalised trust*.

Kenneth Newton's (2007; Newton & Zmerli, 2011) overview on trust emphasises that there are fundamental differences between the two types. *Particularized* trust—or thick trust, as coined by Putnam—is the type of trust that emerges in relatively small communities with high levels of personal interaction amongst all its members. Such communities usually generate high levels of particularised trust due to the common interests and preferences of their members. Perhaps the most important factor is that individuals have personal relationships, meaning that trust is bound to persons more than it is to a concept (Newton, 2007). However, the high levels of ingroup trust often manifest such that outsiders are distrusted. As trust is built upon common characteristics and personal relationships, it does not easily translate onto outgroups and individuals outside the community. They are more likely to be distrusted due to how particularized trust is generated. Trusting that your neighbour of ten years will not steal your belongings while you are at work is easy. Trusting that a wayfarer will act in the same manner, much less so. The mechanism of how particularised trust develops does not easily lend itself to those deemed as "others".

Yet, to varying degrees, people do have some confidence that even the wayfarer will not rob them blind whilst they are away at work. This concept is known as *generalised trust*, the type of trust that exists in greater society and builds less on personal interaction but rather social-institutional[1] and attitudinal trust towards members in society that you don't personally know. Whereas the sources of particularised trust generation are intuitively understandable, generalised trust is a far less tangible concept. Where particularised trust must largely be understood at the individual and community level, and political trust more at the societal level, generalised social trust exists somewhere in the middle, as depicted in figure 1. It does not entirely depend on personal experience with specific individuals, nor does it represent one's trust in an abstract, faceless, institution.

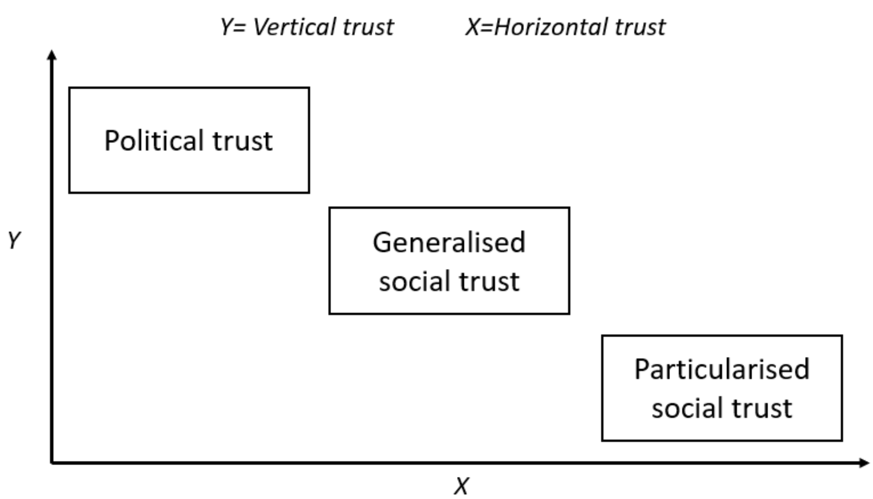

Fig. 1. Own figure. Types of trust plotted in terms of horizontal and vertical influence.

[1] Institutional is in this regard a broad definition of cultural norms, rules, and practices in society where trust can be understood to exist despite distanciation in either time or space where the full information on the social phenomena are not known (Ritzer & Stepnisky, 2014, p. 545)
opposed to state institutional trust that is often used vis-à-vis political trust and encompasses the state's formal institutions.

Generalised social trust can best be understood as something influenced by the general state of things, whilst also being shaped by personal interactions. The intuitive understanding of generalised trust is simple, yet meaningful: how much do you trust people? Untangling what it precisely is, not so much. This is reflected in the literature, and the common consensus is that despite its difficulty in pinpointing, one can measure it with a phrasing much like this: To what extent can you trust strangers? (Bjørnskov, 2012; Newton, 2007). Strangers are inherently different from people you know, as one cannot have personal relationships with a stranger, thus effectively separating generalised trust from particularised trust as a concept. This is important to keep in mind when turning to the empirical literature, and specifically for African countries, as different studies investigate different types of trust.

2.3 Why Social Trust Matters

Fukuyama aptly states that "as a personal attribute, trust is not inherently good or bad" (2015, p. 123). Being a high-trusting individual in a society with high levels of crime and swindling will lead to that individual being taken advantage of. As an attribute of a *society*, however, trust brings a lot of benefits. Economically, it lowers the transaction costs for conducting small-scale trading and improve investment and growth (Zak & Knack, 2001). If you trust those you are engaging with economically, less time and resources are needed to do background checks and ensure that contracts are watertight. This is supported by Bjørnskov (2012), who also comments that numerous studies "have found trust to be an important and robust determinant of economic growth" (2012, p. 1359). Economic cooperation between individuals is, simply put, easier in high-trust societies.

In terms of social benefits, the advantages of a high-trust society are several. When living in a low-trust society, time and money must be invested in ensuring one's protection of both person and property. If you do not trust the people around you to leave your home or business unattended, it puts severe limitations on what you can devout time and resources to otherwise do—such as engaging in community and civic affairs. The much cited article by Lewis & Weigert (1985) goes so far as to state that "Indeed, if trust declines below the barest acceptable level, generalized exchange networks eventually collapse completely" (Lewis & Weigert, 1985). Living in a high-trust society mitigates transaction costs economically, whilst trust in fellow citizens enables social cooperation and coordination whilst minimizing the risks of doing so.

Engaging in community and civic affairs becomes easier, and in turn lays the groundwork for the social institutions that nurture democracy (Newton & Zmerli, 2011; Rothstein & Uslaner, 2005; Uslaner, 2002; Zmerli & Newton, 2008).

Fukuyama (2015) argues that generalised trust seems to be necessary for a well-functioning civil society. Yet, constructing a well-functioning society appears problematic or outright impossible without first a solid basis of generalised trust. This can be understood akin to the security dilemma: trusting that others have good intentions may prove costly, whereas distrusting others ensure one's safety. The problem—like the security dilemma—is that others will respond with distrust as well, regardless of their intentions, resulting in lower generalised trust. This becomes a vicious cycle, and it ultimately results in a collective action problem (Fukuyama, 2015, p. 124). In these cases, Fukuyama, writing on the case of the Italy from the 1860s, observes that people revert to their close community and nuclear family for sources of trust—developing particularised trust. This in turn further feeds into the distrust towards greater society.

The logic of spiralling distrust can easily be understood. How high levels of trust both develop and can be constructed into a virtuous cycle, is harder. Freitag & Traunmüller (2009) points to two theoretical arguments on the formation of social trust: the first posits that trust is based on a form of evaluation of one's environment and is foremost based on prior experiences. In this tradition, which is largely based on rational-choice theory, a positive change in social trust is indeed a collective action problem, but nonetheless one that can be overcome: trustworthy behaviour induces trustworthy behaviour. On the other side, Uslaner argues in *The Moral Foundations of Trust* that "trust must be learned, not earned" (2002, p. 77). Uslaner's position can be placed within what Freitag & Traunmüller (2009) labels *personal predisposition*, or the psychological perspective. In this tradition, it is the world-view and personality of each person that indicates their trust towards strangers, and ought not be prone to change (Freitag & Traunmüller, 2009).

Instead of choosing one tradition over the other, my argument aligns with the empirical findings of Freitag and Traunmüller (2009) by acknowledging that both perspectives possess intrinsic strengths. My view can be summarised in that I assume social trust to be influenced by personal predisposition, and in larger society therefore to be rather stable. However, large changes in one's social surroundings—such as a change in the majority ethnic group of refugees as this thesis analyses—can be substantial enough to warrant adjustment of one's attitudes of trust towards others.

2.5. Ethnicity

Social trust is a complex concept, yet there is a relative consensus of how it should be measured, that it is good thing, and that one arguably cannot have too much of it. Ethnicity is more difficult. Chandra (2006) notes that there has been a general convergence towards which identifiers should register as ethnic by comparative political scientists, yet central scholars utilise definitions that do not match this classification (Chandra, 2006). Her remarks are illustrative of the overall problems of research dealing with ethnicity as a variable. Similar to trust, most people have a general idea of what ethnicity is, yet pinpointing what ethnicity is, is difficult. This has led to a rich debate on what ethnicity is; how it is measured; what separates the concepts of identity, nationality, and ethnicity; and even if ethnicity is a naturally existing or can be a social or political product (Chandra, 2006; Turton, 1997).

Chandra is foremost determined to define ethnicity as a classification of the term for comparative political scientists, not to define necessarily an objective er even day-to-day definition. She defines: "Ethnic identity categories, I propose, are a subset of identity categories in which eligibility for membership is determined by descent-based attributes" (Chandra, 2006, p. 400). Her definition offers an important distinction between ethnicity and identity. These two related but ultimately different concepts are often conflated by the common characteristics described to them both. These include commonness in language, culture, homeland, ancestry and so forth. What then, is ethnicity more specifically? Chandra argues that there are two properties that are intrinsically bound to ethnicity, but not identity: foremost visual, descent-based attributes that are genetically acquired through a cultural and historical inheritance such as names, place of birth, and origins of one's ancestors. In a nutshell: if a person is eligible for membership in an ethnic identity category, then so is automatically their siblings (Chandra, 2006). Whereas identity is fluent and prone to change—e.g., an immigrant integrating and taking on the national identity of the host-community—ethnicity is "sticky" (Chandra, 2006). Sticky, as in that the characteristics of ethnicity cannot easily be changed.

Chandra's definition provides a necessary precision for study, yet it also narrows the scope of what ethnicity and ethnic identity categories *may* represent. Rüegger (2019) defines ethnic groups as “self-perceived communities with a shared culture and a common ancestry” (Rüegger, 2019, p. 45). It stands as a vaguer definition not necessarily differentiated from identity, but it highlights an important element: *self-perceived*. Perceptions of both one’s own—and perhaps as importantly, others—ethnic group is a vital component of analysing how change of the refugee ethnic groups may moderate the impact of hosting refugees. Precisely *perceptions* of ethnic identity, differences, and group membership is what would account for changing attitudes of trust, not ethnicity in itself. I therefore frame the conceptualisation of *why* ethnicity matters in terms of group membership.

2.4 Migration and refugees

A few remarks concerning migration and refugees are also in order. Whereas migrants are associated with a sense of wilful movement, be it for education, work, or otherwise, refugees are people who unwillingly have fled their country and been recognised as a refugee. The 1951 Convention relating to the Status of Refugees defines a refugee as “someone who is unable or unwilling to return to their country of origin owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, or political opinion” (UN, 1951; UNHCR, 2010a). An asylum seeker on the other hand, is someone claiming refugee status but has yet to receive it. In both datasets and assessment of settlement populations, this thesis exclusively considers refugees.

Another noteworthy distinction is between internally displaced persons (IDPs) and refugees. Any conflict is likely to force numerous people to flee their homes, but only those moving across a state border and is recognized as a refugee, are, by definition, a refugee. As I will analyse changes in the refugee group in the host-community neighbouring a country with a conflict, only refugees, not IDPs, will be part of the analysis.

3. Literature Review

The three fields in the previous chapter all have rich literatures to draw from. There is, however, limited research specifically for the relationship between social trust, refugees, and ethnicity[2].

[2] A substantial literature has developed on the effects of hosting large refugee populations in Sub-Saharan Africa that is not treated here due to parsimony. Noteworthy is adverse economic effects (Alix-Garcia & Saah, 2010;

Foremost studies that lends insight into how trust is affected by either ethnicity or refugees are of central interest.

Trust as a topic of study has received much scientific interest since the 1990s and its importance has become widely accepted (Freitag & Traunmüller, 2009; Rothstein & Uslaner, 2005). Despite this, studies on how refugees impact perceptions of trust in the host community are largely confined to European experiences and survey data. One such example is McLaren's (2012) study on political trust in Great Britain, finding "that concerns about the impact of immigration significantly [negatively] affect *political* trust" (McLaren, 2012, p. 163). Her 2017 study expands to countries in West and Southern Europe, largely confirming her prior findings. A different but complementary view is presented in Herreros & Criado (2009) that finds that socially high-trusting individuals have more positive attitudes towards immigrants.

Dinesen & Sønderskov (2015) more closely investigate the relationship between social trust and ethnicity on the micro-scale in Denmark. Operating with data able to identify the ethnic diversity in a radius of 80 meters of a respondent adress, they establish a clear negative link between ethnic diversity and social trust. However, when aggregating the data into larger contexts, this effect gradually vanishes at measurement levels above 180-250 meters. This gives rise to their conclusion that the underlying mechanism of negative social trust in this context is exposure.

Robert Putnam's (2007) article on ethnicity and social trust in the USA reports similar findings. In ethnically more diverse neighbourhoods trust is lower both towards ethnic in- and outgroups for the respondent. This is supported in Stolle et al. (2008) finding a similar effect with white majorities. Putnam further argues that in ethnically diverse neighbourhoods, all the inhabitants "hunker down": in various forms retreating more from public life than they do in ethnically less diverse ones. A view also empirically supported in Alesina & Ferrara (2000). A recent contribution by Lundåsen (2023) present a nuancing picture. She finds a positive moderating effect of civil society engagement on generalised social trust in Sweden where asylum-seekers are more numerous, as opposed to those who do not engage in such activities. Her findings are strikingly similar when outgroup trust is replaced as dependent variable instead of social trust.

Callamard, 1994; Chambers, 1986), but also positive economic effects (Jacobsen, 2002; Kreibaum, 2016; Maystadt & Duranton, 2014; Maystadt & Verwimp, 2014; Taylor et al., 2016).

In varying ways, these studies provide some evidence that ethnicity and perceptions and attitudes towards it, affect trust *in general.* Yet, what these studies all share is that they exclusively investigate Western countries, i.e., the Global North. Particularly European countries constitute fairly ethnically homogenous populations. Putnam's study on the US sets it apart in this regard, yet the US shares with the European countries virtually every other measure. Western respondents are, on a global scale, not particularly representative: they generally have high living standards, high levels of education, have country borders that correspond with a national identity, and live in societies where the state is efficient and supports strong institutions and economies. In other words, these are societies where one would expect trust to flourish (Fukuyama, 2015). The Global South, and Africa especially, is far more diverse in all these areas, and as expected, also has far lower levels of trust (Logan & Torsu, 2022).

Delhey & Newton (2005) study the conditions that are associated with high social trust. They study the percentage of a country population who has stated they trust others on a dichotomous variable. Their results largely confirm the global North-South divide with a sample of 60 countries. They find that high trust countries usually are characterized by ethnic homogeneity, Protestantism, wealth, and good governance. The authors conclude with the question as to how generalised generalised social trust really is, when ethnic homogeneity evidently is that important for social trust.

Thus, the idea that European insights are readily exportable is problematic, as has been shown by several studies. E.g., Güemes & Herreros (2019) compare the effect of education on generalised social trust in Europe, Latin America, and Africa. Whereas education has a strong positive effect on trust in Europe, and a weaker yet positive effect in Latin America, the effect is strongly negative in Africa. They conclude this to be a trust paradox: in states with efficacious institutions where trust is less needed, it is abundant. In failed and weak states with far less efficient institutions, social trust is sorely lacking.

Dinesen (2012) therefore presents a particularly interesting study. He flips these insights and asks if (dis)trust travels by analysing immigrants in Northern Europe's social trust. His findings indicate that immigrants adapt to the hight trust levels witnessed in the host-community to a great extent. He notes that this is contrary to findings in the US (Uslaner, 2008), and presents a possible explanation to be ethnic heterogeneity (Dinesen, 2012). Such an explanation is examined in the review study by Meer & Tolsma (2014). Synthesising 90 studies on ethnic diversity and social cohesion, they find that there is a negative effect when the object of study are neighbourhoods. They furthermore assess that “people in ethnically heterogeneous environments are less likely to trust their neighbours or to have contact with them. However, this does not spill over to generalized trust” (Meer & Tolsma, 2014). The studies included are nonetheless mainly on the Global North. What then, of the relationship between ethnicity, social trust, and refugees in Sub-Saharan Africa, that differs so starkly in all three categories?

3.1. Ethnicity, social trust, and refugees in sub-Saharan Africa

The study that most closely investigates the mechanisms of ethnicity and social trust in Africa is Hodler, Srisuma, Vesperoni & Zurlinden (2020). They construct an ethnic stratification index comprised of the extent socio-economic hierarchy exists along ethnolinguistic lines in 26 African countries. Rothstein & Uslaner (2005) have likewise presented the argument that equality is an important determiner of social trust. The key finding of Hodler et al. (2020) is that where ethnic stratification is high, trust in relatives, neighbours, and acquaintances is significantly lower than in societies with low ethnic stratification. This is however a measurement of particularised social trust. As mentioned above, particularised and generalised trust are conceptually different, and the mechanisms of how they develop differ.

Moscona, Nunn & Robinson (2017) study social trust, and in Sub-Saharan Africa. Unlike Hodler et al., they measure both generalised and particularised social trust. They analyse how individuals in segmentary lineages—that is, more traditional societal structures such as tribal organisation—differ from others in the levels of trust towards relatives and nonrelatives. Their central finding is that there is a substantial gap between the two varieties of trust. Individuals in segmentary lineages display lower trust in nonrelatives, and they further hold that this is causally not due to their relative higher trust in relatives. This is as opposed to individuals not living in segmentary societies, where individuals have higher levels of generalised trust. Both Hodler (2020) et al. and Moscona et al. (2017) provide empirical evidence that the structure of society and ethnic composition affect how different varieties of social trust develops, manifest, and differ. In other words, ethnic differences can be viewed as a barrier towards the development of generalised trust.

The arrival of refugees—and the subsequent altering of the ethnic composition of society—may then have a series of effects. One such effect is the likelihood of conflict, as investigated by Bertinelli, Comertpay & Maystadt (2022). A substantial literature with mixed empirical evidence points to refugees being associated with higher conflict risk (Zhou & Shaver, 2021). Bertinelli et al. (2022) introduce changing ethnic composition as an intermediary variable in explaining this link. They analyse the likelihood of conflict with changes in ethnic diversity and polarisation in refugee-hosting areas. They find a positive relationship between refugee arrivals and conflict, however, emphasising that this is not due to refugees per se, but rather that conflict likelihood increases when refugee arrivals exacerbate already existing ethnic polarization. Yet, where refugees increase ethnic *fractionalisation*, the risk of violence is lessened. These findings run in contrast to the recent development in migration and conflict studies, where e.g. Zhou & Shaver (2021) find no evidence that hosting refugees are linked to either prolonging conflicts or increasing the likelihood of new conflict or violence. Although the empirical evidence is contested, Bertinelli et al. (2022) show the importance of analysing changes in ethnic composition as an intermediary explanation of the effects of refugees. It is precisely such change that is analysed in this thesis.

This importance is further actualised by Rüegger & Bohnet's study (2018) on refugee flight patterns. By analysing yearly outflows of refugees and controlling for a series of pull factors, they find that where refugees flee is not random. Rather, it is largely determined by ethnic kinship in neighbouring countries and whether the prospected host country has a record of being accepting towards co-ethnic refugees in the past. Despite the problems of acquiring accurate data on refugee ethnicity, they report that between 1975 and 2009, 80% of refugees relocated to a neighbouring country and that 46% of refugees had ethnic kin in the country of asylum (Rüegger & Bohnet, 2018). Their analyses further show that "refugees consider cultural pull factors directing them towards certain countries of asylum" (Rüegger & Bohnet, 2018, p. 82). Moreover, they find that the larger the co-ethnic group in the potential host country, the larger the predicted count of those refugees; that the relative capacity of the co-ethics in the host country matters; and that refugees often follow the patterns of prior migration groups due to established transportation networks and aid facilities (Rüegger & Bohnet, 2018). In other words, where refugees flee are highly dependent on ethnic kinship. For the purposes of this thesis, their findings as well as their dataset is used to contextualise and frame the case selection.

4. Theoretical framework, assumptions, and hypotheses

Combined, the conceptual framework and previous studies showed some evidence that (i) generalised and social trust differs between African societies of different organisation; (ii) that ethnicity and ethnic differences have an effect on how the two types of trust differs; (iii) that refugee arrivals can exacerbate and influence ethnically linked challenges in the host community; and (iv) that where refugees flee heavily relies on the existence of co-ethnics in the host-community.

Based on these previous findings, I propose the argument that refugees will most likely influence the ethnic composition in the country of arrival and could exacerbate ethnic divisions, further entrenching the ethnic barrier to the development of generalised social trust.

As discussed in the sections of social trust, the mechanisms of ingroup trust generation does not extend well to the outgroup. What constitutes a "group" in this regard? I propose that ethnicity is a good candidate: "in many societies ethnicity, real or imagined, is the basis for social and political identity" (Collier et al., 2001). However, other forms of in- and outgroups can exist. The empirical evidence from the European experience indicates that refugees—regardless of ethnicity—is associated with a decrease in social trust. Without taking ethnic *linkages* into account, one can theorise that the relationship is similar in an African context. Refugees are often in precarious situations. Lack of food, proper shelter, and the general means of survival can force people into extreme actions. Refugees can in this regard be viewed as an element of uncertainty, giving rise to more protective attitudes in the host community. Refugees, on the merits of being refugees as opposed to the members of the host-community, can be viewed as an outgroup in this regard. This narrow understanding gives rise to H1:

H1: Refugees negatively impact the level of generalised social trust.

H1 can be seen as a baseline hypothesis where ethnicity is less important. Keeping ethnicity out of the mix is nonetheless only theoretically possible, as a plethora of the studies highlight ethnic diversity as a component negatively influencing social trust. In- and outgroup and ethnic diversity in a European context is usually inseparable from arrival or presence of refugees. The reason being that European nation-states in general constitute quite homogenous societies, and the newly arrived immigrants become a clear out-group[3].

Turning to the African context, would refugees present an ethnic outgroup? Not necessarily, as state borders seldom correspond to ethnic, religious, or historical borders (Wishman & Butcher, 2022). The research by Rüegger & Bohnet (2018) shows that refugee flight patterns are not random, but rather that refugees flee to countries that have ethnic kin populations, i.e., refugees should not necessarily be interpreted to be an outgroup in the receiving country.

Depending on the ethnic linkages that exists across borders, refugees will become part of some host community members ingroup, and others' outgroup. Depending on which one of these ethnic groups are the most prevalent in the host community, generalised social trust may be impacted differently. The mechanism of trust generation in relations to groups lends to the hypothesising that if refugees share ethnic linkages, generalised social trust ought not to become lower for the host population. If the opposite is true, out-group distrust and strong intergroup particularised social trust may be exacerbated by the arrival of ethnically out-group refugees, leading to lower generalised social trust in the host community. The reasoning is that the probability of developing trust with people outside your group decreases with the social distances between one's group and out-groups (Finseraas & Jakobsson, 2012). I therefore theorize that H1 will be heavily moderated by ethnic differences between the refugees and host-community members.

H2: The negative effect of refugees on generalised social trust in the host community is moderated by ethnic differences.

Differences can be understood and analysed in a wide variety of ways. The next chapter on research design describes the reasoning and benefits of operationalising *differences* as *change* in the majority ethnic groups.

[3] An important addendum to this is the Ukrainian refugee crisis, where European attitudes towards refugees arguably differs quite a lot from the attitudes during the 2015 refugee crisis.

5. Research design and case selection

The hypotheses are fairly straightforward, but adequately testing them is not. The data and methods section will greatly expand on the research design, but a few remarks are necessary as it frames the entirety of the thesis. I analyse individual level generalised social trust in the refugees' host community with Afrobarometer data to ascertain how refuges impact social trust. The refugees themselves will therefore not be analysed, only host-community respondents.

To ascertain more than mere association between refugees and social trust in the host community, it is crucial to establish measurements of social trust before and after arrival of refugees. If refugees affect social trust in the host-community, one will only expect this to be evident where the refugees reside, not elsewhere. Spatially isolating the respondents that live in proximity to refugees to those who do not, enables a far more precise analysis.

This translates to a difference-in-differences design, or before-and-after comparison, as it is commonly referred to. Difference-in-differences (henceforth DD) is a form of quasi-experimental design. -Experimental as in it relies on a separation of control and treatment group, as well as before and after measures. Quasi-, as in that it is not a *true* experiment which is often viewed as the scientific golden standard. Unlike the true experiment where the control and treatment group can be chosen and manipulated to the researcher's whims and wishes, the quasi-experimental researcher relies on finding suitable control and treatment groups out in the wild world of available data. Unfortunately, it is in the quasi-experimental design's nature that the perfect treatment and control groups do not exist.

The core reasoning behind DD is to (partly) overcome this problem. DD assumes that the control and treatment group will have other factors that influence both groups differently before the treatment (Gertler et al., 2016). I say partly because it is further assumed that if the change did not occur for the treatment group, it would develop in the same manner as the control group, known as the "equal trends" assumption (Gertler et al., 2016).

Finding a control group (no refugee exposure at both points of measure) and a treatment group (refugee exposure only after the first point of measurement) that fulfil this criterion is central. This is, however, only half the question, as my hypothesis posits that ethnicity will moderate the effect. Some remarks concerning operationalisation of *ethnic differences* is necessary to illustrate the different treatment and control groups I seek in the natural experiment. To be able to analyse ethnic difference, a third respondent group needs to be accounted for, namely a group which is exposed to refugees at the first measurement but exposed to a different ethnic group of refugees for the second. This is what gives rise to the unique triple difference estimator (DDD), greatly expanded on in the methods chapter. Consider the groups necessary in table 1.

Table 1: DDD respondent groups

	First measurement	Second measurement
No refugee exposure group	No exposure	No exposure
Refugee exposure group	No exposure	Refugee exposure
Ethnic change in the refugee population group	Refugee exposure	Ethnic change in the refugee population

A common way of analysing the effects of ethnicity is by utilising heterogeneity/homogeneity or fractionalisation indices (as used in e.g. Denny & Walter, 2014; Fearon & Laitin, 2003; Knack & Keefer, 1997). Posner (2004) makes a compelling argument against such indices precisely because they do not account for spatial and relative distribution of groups. Quantifying a measure of *ethnic differences* between refugees and host-community is, however, both difficult due to data limitations, as well as being inherently problematic and ill-advised. Another option is ethno-linguistic distance, as e.g. Hodler et al. (2020) utilises. However, as Chandra's (2006) discussion highlights, language does not equate ethnicity. E.g., Hutus and Tutsis have virtually no linguistic differences (Huening, 2013), yet constitute perhaps the most infamous ethnic conflict in Africa.

Instead of acquiring erroneous measurements of ethnic differences, fractionalisation, or some computed "ethnic distance"—however that might be conceptualised—I will base my analysis on an exogenous variable of ethnicity: change. Change, as in that the majority ethnic group of the refugee population changes between the measurements of trust in the host community. In this way, regardless of whether the ethnicity of the refugees translates to either an in- or outgroup for the majority of the host-community, the change in the refugee group presents a break from being either in- *or* outgroup for the majority of the host-community respondents. Such an operationalisation also takes *perceptions* of ethnicity into account.

It therefore becomes necessary to identify a specific subgroup of host-community respondents that experience a change in the ethnic composition of the refugee population between the points of measurement.

The subsequent sections of case selection are the precise endeavour of “finding” a natural experiment in Sub-Saharan Africa where these respondent groups can be identified. To do this, quite an expansive list of information needs to be present for the cases. Particularly difficult is obtaining any form of ethnic data on refugees. The proper presentation and treatment of the data and the problems they solve—as well as pose—will be outlined in the Chapter 8. It is however important to note that the pre- and post-measures cannot be changed. Generalised social trust is only measured in the geocoded Afrobarometer survey rounds 3 (2005/2006) and 5 (2012/2013).

“Finding” cases with the first and second groups in Table 1—the groups where ethnicity is considered irrelevant—is the least problematic. Refugee flows are large, varied, and nearly constant between a great deal of Sub-Saharan African countries. Finding such cases where the ethnic group of refugees also changes is far more difficult. I have therefore pursued finding such cases.

To overcome this empirical challenge, I have chosen to fix these elements around a specific conflict. This offers numerous advantages in case selection: (i) it ensures that the ethnically different refugees are arriving after the first measurement of social trust, and (ii) By choosing a conflict that erupts between the two measurement periods, i.e., two Afrobarometer rounds, the country of origin for the refugees remains constant for a subgroup of refugees in the cases to be analysed, greatly alleviating the problems of correctly identifying the ethnicity of the refugees in a given refugee settlement.

6. Case selection: Framing conflict

The choice of countries to be included in the analysis is thus dependent on an external conflict in a country that will not be part of the analysis. The choice of the framing conflict, or refugee-generating event, is therefore pivotal for the analysis. Interestingly, 2005-2013 corresponds to the most peaceful period in post-colonial Africa, as depicted in Figure 2.

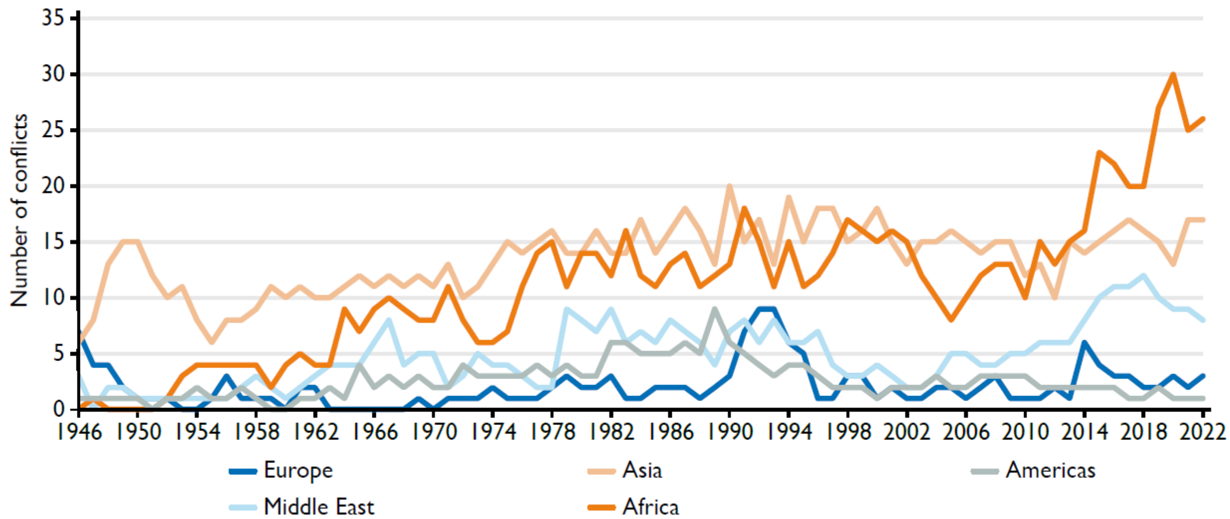

Fig. 2. From Obermeier & Rustad (2023). Number of state-based armed conflicts by region, 1946-2020. Source: UCDP/PRIO Armed Conflict Dataset (Davies et al., forthcoming).

That is not to say that there were no conflicts that erupted between 2005 and 2013, many did but were not chosen based on three criteria: (i) The conflict must be of a magnitude to create a noticeable increase of refugees to neighbouring countries; (ii) the country of conflict must have neighbouring countries with overlapping data for both 2005 and 2013; and (iii) the conflict must develop in a way that makes a change in the refugee group likely.

Several conflicts were considered but abandoned: The Boko Haram insurgency in Nigeria, due to spread and religious overtones. In the case of the Chadian Civil War, only survey data exists for Cameroon, leaving the analysis with a singular case. This was the criterion that most frequently caused a conflict to be abandoned as a framing event for the choice of countries to be analysed. Figure 3 presents a map for overlap between refugee settlement (Fisk, 2019) and Afrobarometer respondents.

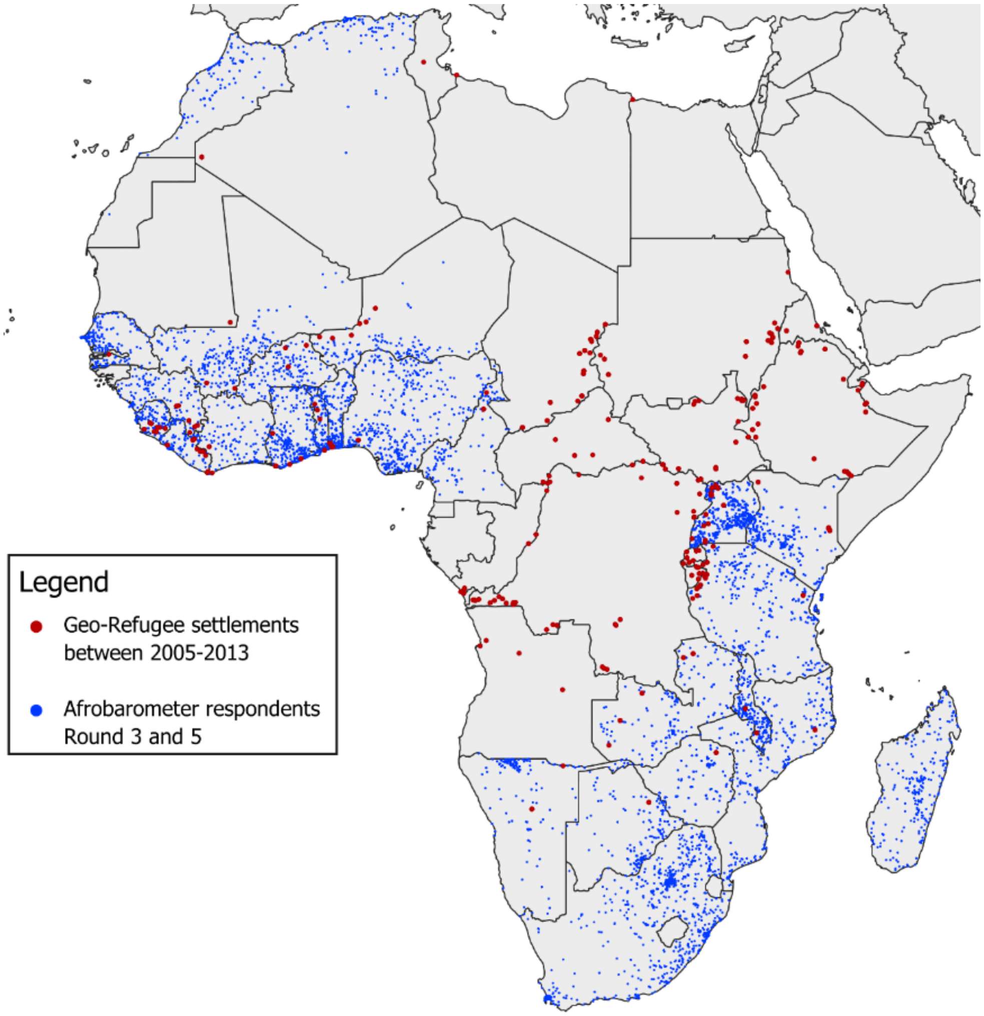

Figure 3. My visualisation using QGIS software. Map of Africa with Afrobarometer respondents in round 3 and 5 and UNHCR settlements 2005-2013 from Geo-Refugee dataset (Fisk, 2014)

6.1 Kivu Conflict

Based on these considerations, I chose the Kivu conflict in the Democratic Republic of the Congo (Henceforth DRC) as the framing conflict. Like the Chadian civil war, the conflict is mainly confined to the border region, increasing the likelihood of people fleeing the country rather than becoming IDPs. Figure 4 shows a map highlighting the Kivu region, the borders to Uganda, Rwanda, Burundi, and Tanzania as well as an overlay of UCDP fatalities between 2006-2013. Unfortunately, neither Rwanda nor Burundi participates in Afrobarometer rounds 3 and 5. The cases, as I will return to in the next section, will therefore be the Afrobarometer participating neighbouring countries Uganda, Tanzania, and Zambia.

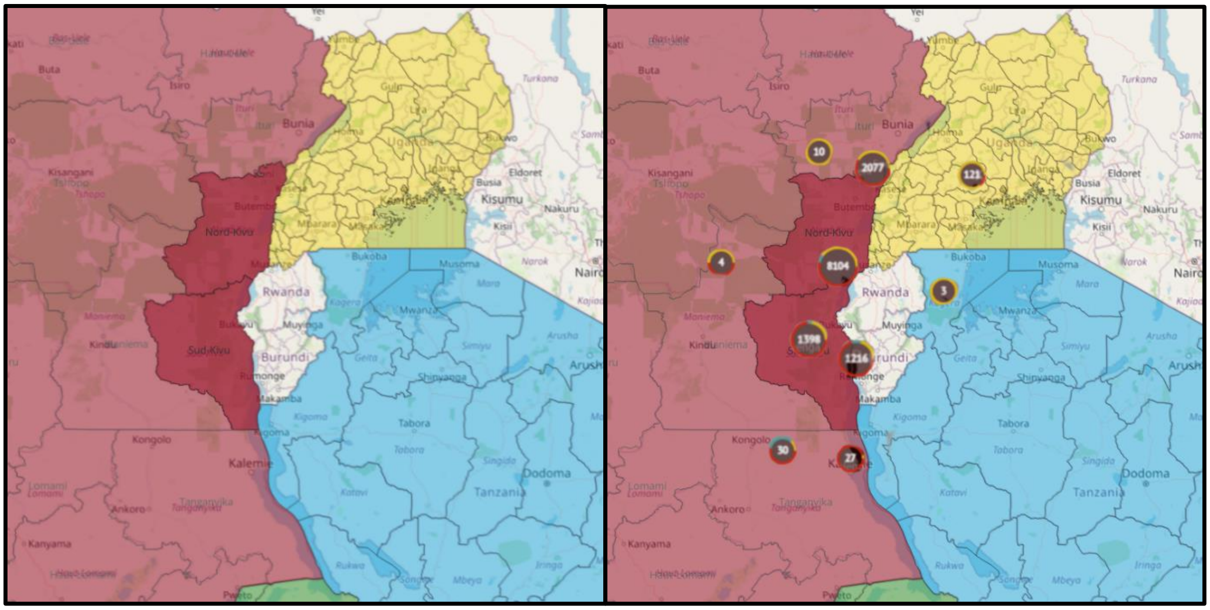

Figure 4. Left: Map of Eastern Congo (light red). The Kivu region (North and South Kivu) is highlighted in dark red. Uganda, Tanzania and Zambia are outlined in yellow, light blue, and green, respectively, as they will serve as cases. Map made in QGIS: GADM 4.1 shapefile and OpenStreetMap (OSM). Right: Overlay with UCDP Fatalities between 2006-2013: https://ucdp.uu.se/exploratory

The Kivu conflict arguably began as early as 2004 when General Laurent Nkunda rebelled, a year prior to the first measurement of social trust. Yet, the fighting did not escalate until 2006, and the UCDP/PRIO Armed Conflict Dataset version 22.1 specifies the start date as 01.11.2005 (Davies et al., 2022; Gleditsch et al., 2002). I will not devote a large chapter to the specifics of the conflict. However, as Jason Stearns, former Coordinator of the UN Group of Experts on the DRC states: “The continuing violence in the eastern Democratic Republic of the Congo (DRC) can be bewildering in its complexity: in the profusion of armed factions, the plethora of acronyms, and the multitude of grievances that trigger outbreaks of conflict” (Stearns, 2012). Some mildly confusing background is necessary to understand the change in refugee majority ethnic group.

An important backdrop for the rebellion is the Democratic Liberation Forces of Rwanda (FDLR), a DRC-based group with extremist Hutu ideology that previously has been supported by then DRC president, Joseph Kabila, as a buffer towards Rwanda. The Kivu conflict initially started when the Congolese Tutsi and general in the Armed Forces of the Democratic Republic of Congo, Laurent Nkunda rebelled in North Kivu with the majority of his forces and established the National Congress for the Defence of the People (CNDP). The group and the rebellion were founded with Nkunda "claiming he was trying to stop an ongoing genocide of the [Tutsi] Banyamulenge people" (Uppsala Conflict Data Program, n.d.-b). The CNDP established a state-like apparatus with taxation, administration and police (Uppsala Conflict Data Program, n.d.-b).

Both North and South Kivu have witnessed long and complicated ethnic relations between primarily Hutus and Tutsis. However, the groups have also been lumped together by other Congolese under the colloquial "Rwandophones" (Huening, 2013). The fear of Rwandophones is founded both on conflicting views on origins, particularly regarding the Tutsi-Banyamulenge, as well as a fear of the "Rwandophone rise", the idea that Rwandophones would take over the Kivu region (Huening, 2013). Anti-Tutsi sentiments and the threat posed by the Hutu FDLR are therefore important context for Nkunda's Tutsi rebellion. The rebellion instigated strong anti-Tutsi-rhetoric in DRC, and even direct threats during the 2006 election, capitalising on the growing Tutsi fear (Ochieng, 2017).

The CNDP experienced a split with Nkunda arrested and held by the Rwandan government. With new leadership, the CNDP concluded peace agreements with the DRC government as well as becoming a political party (Uppsala Conflict Data Program, n.d.-a). In 2009 the Hutu extremist group FDLR capitalised on the anti-Tutsi sentiments and begun attacks on Tutsi villages in southern Kivu and reclaiming old territory ('"Dozens Killed" in DR Congo Raids', 2009).

The Tutsi led rebellion and organisation of a state-like apparatus in the early stages translates to the likelihood of few Tutsis fleeing DRC. The subsequent conflict development, where the Tutsi CNDP rebellion is ended and Hutu militias aling with the DRC government, greatly increases the likelihood of Tutsis fleeing DRC. Again, drawing on the findings in Rüegger & Bohnet (2018) one would expect Rwanda to be the natural country of asylum. Unfortunately, neither Rwanda nor Burundi participates in Afrobarometer. Uganda, Tanzania, and Zambia, however, do, and will serve as cases.

6.2 Refugee ethnicity and populations in Uganda, Tanzania, and Zambia

The Kivu conflict as a framing conflict represents two expectations for the cases: (i) that refugee flows from DRC will increase, and therefore also the refugee population in the host-community. This is however not necessarily the case, as domestic refugee policy and refugee return agreements are common. In 2010, the UNHCR published an infographic of the DRC refugee situation, indicating the refugee populations in neighbouring countries; North and South Kivu refugees; as well as returnees, depicted in figure 5 (UNHCR, 2010b).

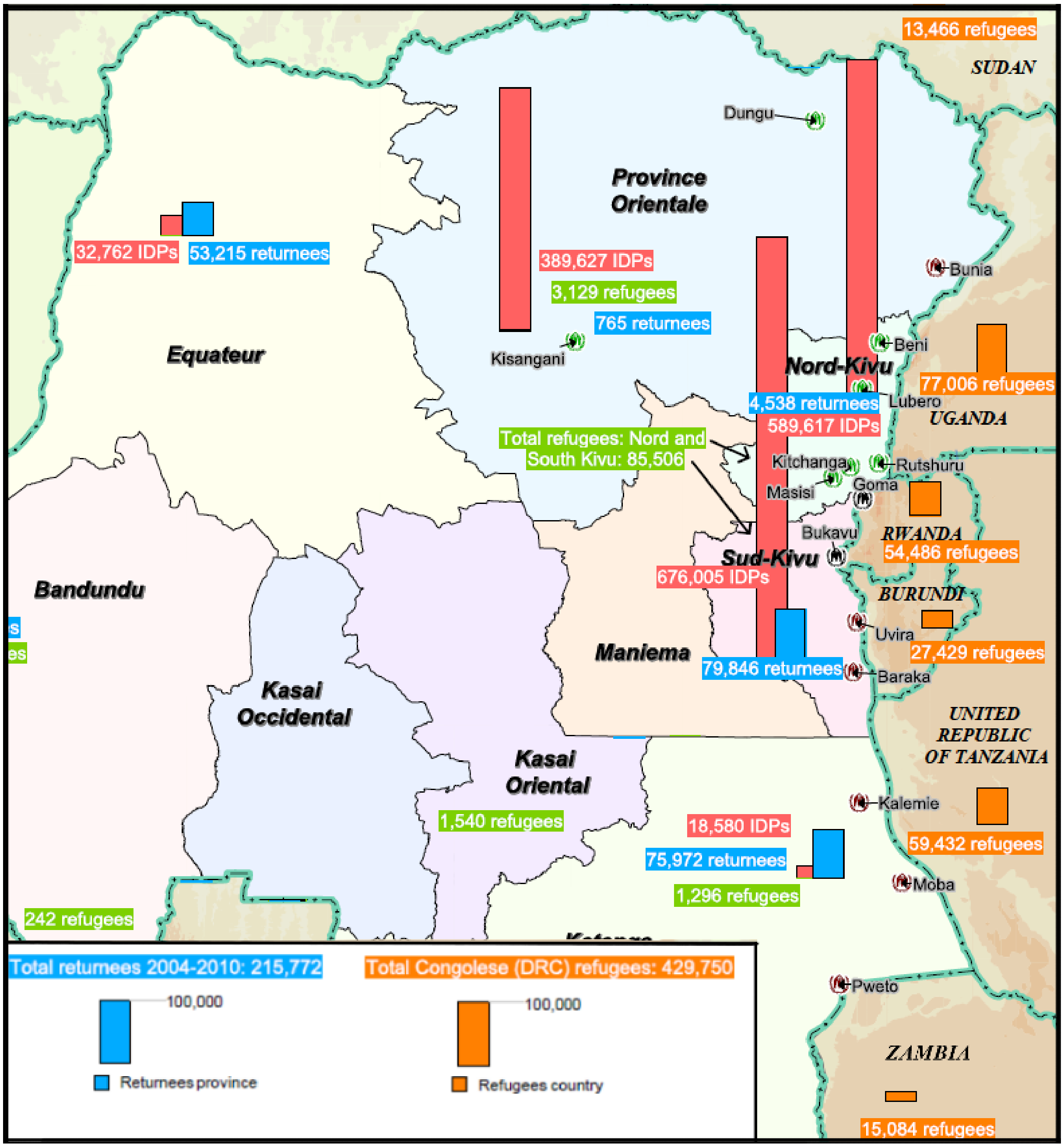

Figure 5. Modified from UNHCR infographic (2010).

The infographic displays a more accurate picture of the Kivu conflict in terms of refugee movements. Note especially that the number of *returnees* is those returning to country of origin between 2004-2010, greatly corresponding to the initial start of the Kivu conflict with the Tutsi-led CNDP's formation. This means that the 85 506 North and South Kivu refugees in 2010 are more likely to be due to the 2009 development. The second expectation (ii) is that the refugees are of a different ethnicity than before, which the development and relative numbers of refugees and 2010 refugees increase the likelihood of being Tutsis.

The infographic does, however, not provide data on such a claim. Only one dataset includes information on the ethnicity of refugees, which is the Ethnicity of Refugees dataset (ER) (Rüegger & Bohnet, 2018). The ER dataset is part of the Ethnic Power Relations family of datasets and defines ethnicity, building on Weber's 1976 definition, "as a subjectively experienced sense of commonality based on a belief in common ancestry and shared culture" (Vogt et al., 2015). This is further specified to include "ethno-linguistic, racial and ethno-religious groups" (Rüegger, 2021). All the following data visualisations are based on the ER dataset.

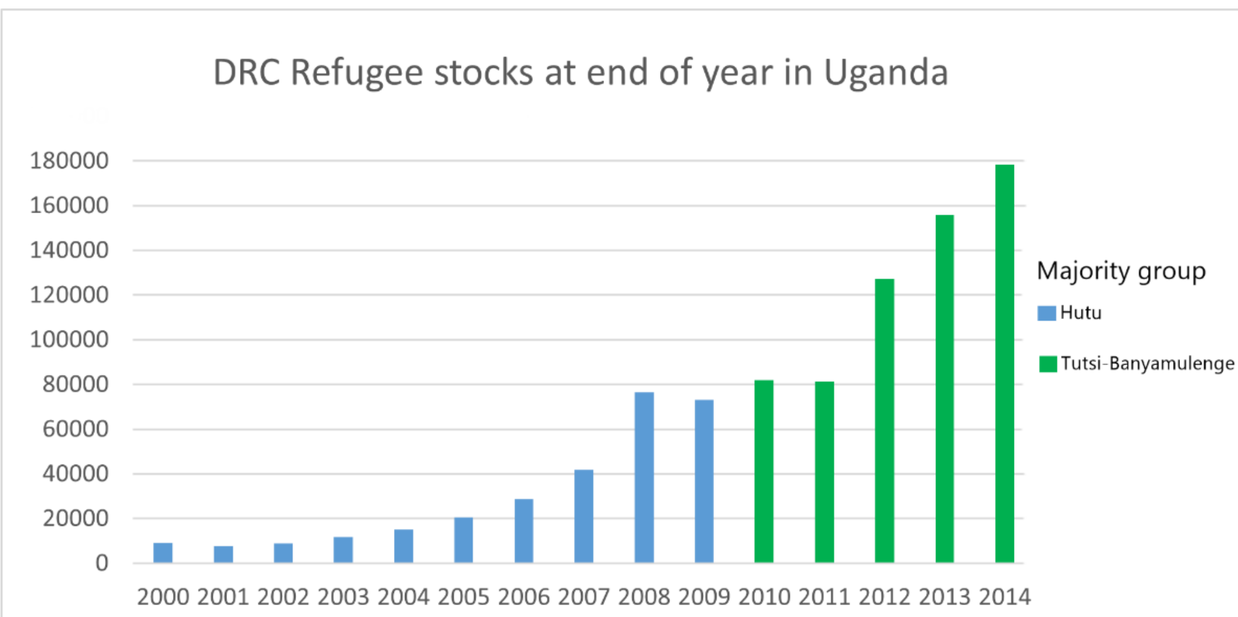

Figure 6. Total DRC refugee populations in Uganda 1975-2013 (Rüegger & Bohnet, 2018). Note: The cut-off is set to 2013 as refugee stocks dramatically increase from 2014-2020, obfuscating the 2004-2012 numbers.

In Uganda, the yearly refugee populations can be viewed in tandem with the outbreak of the Kivu conflict, as depicted in figure 6. The number of DRC refugees residing in Uganda increases from about 20 000 at the outbreak of the Kivu conflict, to close to 180 000 barely nine years later. This very well translates as a population shock. Moreover, the majority ethnic group changes from Hutus to Tutsi-Banyamulenges in 2010.

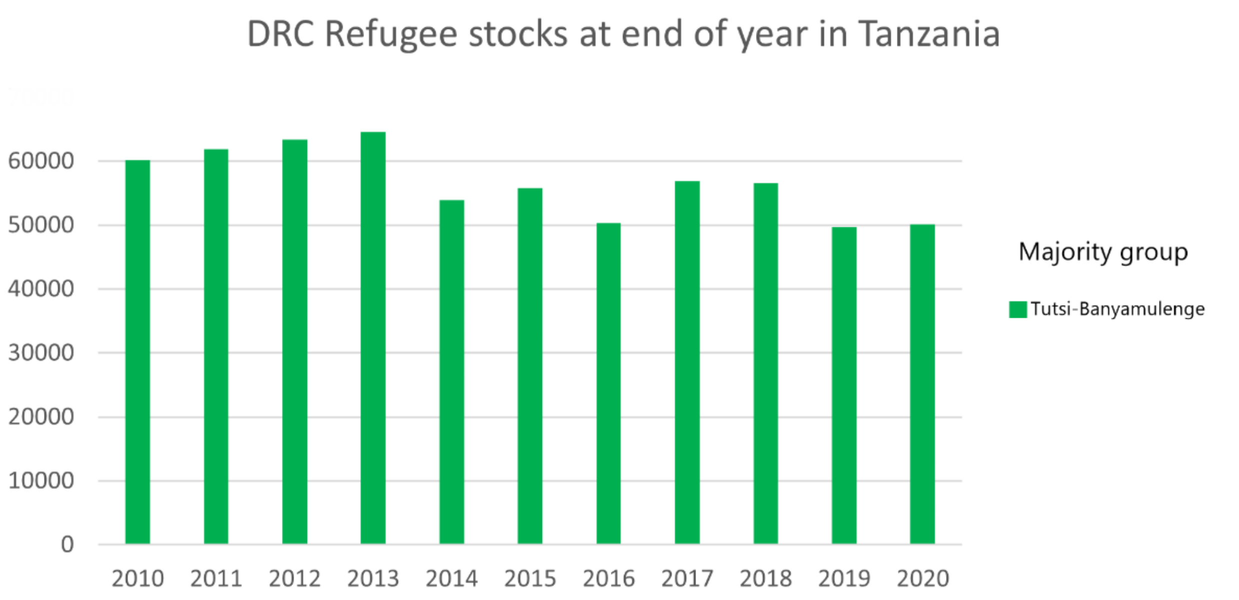

Figure 7. Total DRC refugee populations in Tanzania 2010-2020. Data from the Ethnicity of Refugees dataset and UNHCR (Rüegger & Bohnet, 2018).

Unfortunately, the Ethnicity of Refugees dataset does not provide yearly refugee populations of DRC refugees in Tanzania prior to 2010. This is problematic, as I cannot ascertain that the dominant refugee group changed in 2010, as is the case for Uganda. Due to the relatively fewer DRC refugees arriving in Zambia—seeing that the Zambian border is further away from the Kivu region—I opt to include Tanzania in the analysis regardless. This is the reality of working with difficult data in an African context.

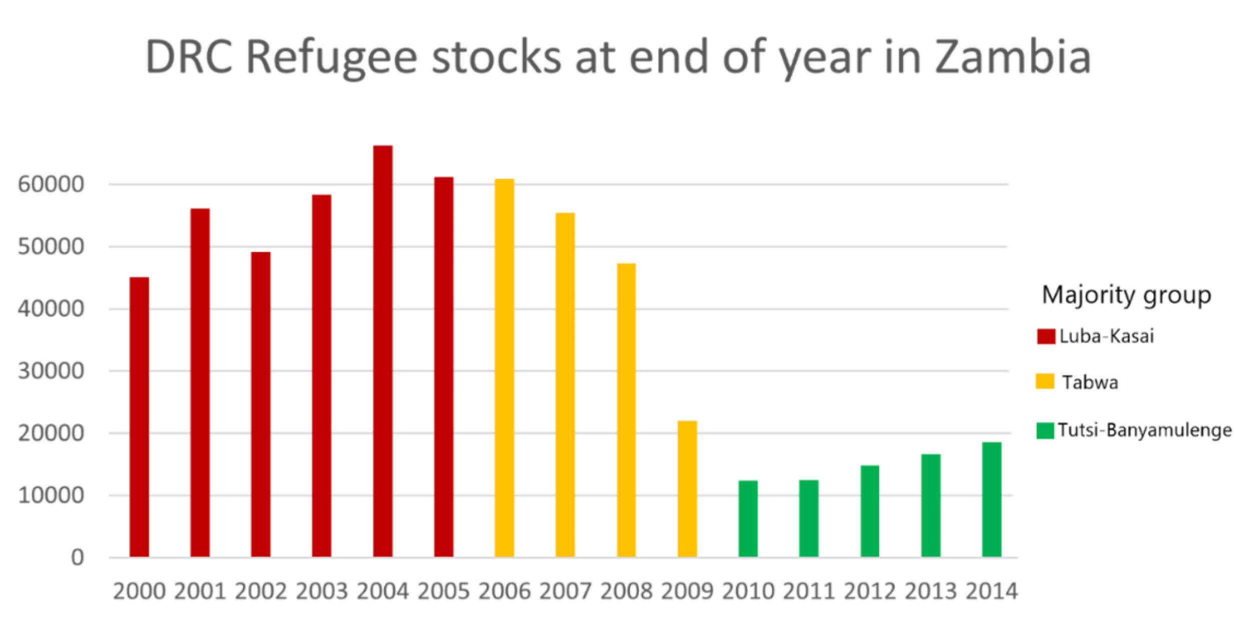

Figure 8. Total DRC refugee population in Zambia, 2010-2020 (*Rüegger & Bohnet, 2018*).

DRC refugee population in Zambia do not resemble the situation in Uganda in terms of numbers. Whereas the Kivu conflict could be argued to create a refugee population shock in Uganda, the opposite seems true in Zambia. However, the dominant DRC refugee group is Tutsi-Banyamulenges from 2010, as is the case with both Uganda and Tanzania.

Combined, the development of the framing conflict, the UNHCR infographic, and the Ethnicity of Refugees dataset provide a solid qualitative and quantitative basis for the core assumption of the thesis:

The major ethnic group of DRC refugees ***changes*** *for Uganda, (Tanzania), and Zambia between the two measurements of social trust, Afrobarometer round 3 and 5.*

DRC refugees only constitute a part of the refugee populations in all three countries. Exposure to the "other" refugees will also be analysed, as a basis to draw comparisons from.

7. Data and variables

With this central assumption made, this chapter is devoted to how I have operationalised treatment variables proxying for refugee exposure. In order for the difference-in-difference methodology to be viable, the variables I construct must separate the host community respondents into a "no settlement group", a "refugee exposure group" and a "DRC refugee exposure group" both temporally and spatially. This chapter is therefore organised around the variables and the various datasets are introduced when relevant. I outline the strengths and problems the various datasets create, and the solutions I have chosen. Seeing that the research question is more complex than available data originally can answer, the data methodology of this thesis is complex. To create reliable and valid treatment variables, I draw on a multitude of methods ranging from the use of GIS software to qualitative assessments of individual refugee settlements to be coded.

Section 7.1 concerns the dependent variable and datasets used. This section also establishes the temporal element of the DD design. Section 7.2 outlines the construction of treatment variables. This is the spatial element of the DD design, as different respondent groups need to be spatially separated to analyse relative change. Given the varying levels of quality in both data availability and the various datasets themselves, substantial work has been needed in terms of data management, sorting, verification, cross-referencing and coding. It follows from this a margin of error. To ensure transparency and reliability of the results, I outline and describe the process in detail. Section 7.3 presents the control variables, and the full sample descriptive statistics are presented in section 7.4.

7.1. Dependent variable: Generalised social trust

The main dataset for analysing African respondents in Tanzania, Uganda, and Zambia is the Afrobarometer dataset (Afrobarometer, 2005, 2015) with geocoding provided by AidData (BenYishay et al., 2017). Afrobarometer is an attitudinal survey of individuals that has been ongoing and developing since 1999.

Afrobarometer is a pan-African, non-partisan organisation that conducts public attitude surveys on democracy, governance, economy, and society financed by a wide range of both government and private organisations, institutes, and centres. It is important to underscore that Afrobarometer is *not* panel data, but cross-sectional. This is not to say that the respondents are unique for each round, but there are no data connecting individuals between different rounds of data collection. I code a round prefix for each respondent to ensure that all respondents are unique.

The dependent variable will be generalised social trust in all analyses. As mentioned, only round 3 and 5 of Afrobarometer includes a variable on generalised social trust, thus establishing the temporal element of the analysis.

Similar to the first question of "the much tried and tested Rosenberg scale" (Newton, 2007, p. 345) the question wording in Afrobarometer round 3 is "Generally speaking, would you say that most people can be trusted or that you must be very careful in dealing with people?" (Carter, 2008). For Afrobarometer round 5, the question wording changes somewhat: "Let's turn to your view on your fellow citizens. Generally speaking, would you say that most people can be trusted or that you must be very careful in dealing with people?" (Park, 2015). This could have a priming effect, in that the round 5 wording specifies fellow citizens, whereas in round 3, "people" have no specification.

For both rounds, this is a dichotomous question, with the answer possibilities of "Must be very careful" and "Most people can be trusted (Carter, 2008; Park, 2015). Arguably, generalised social trust is better analysed as a continuous variable, as pointed out in Newton: "we do not either trust or distrust, but do so to varying degrees. In other words, trust is a variable that ranges along a continuum" (2007). For African survey data, no continuous variable of social trust exists.

7.2. Spatial treatment variables

The major component of the data chapter is the creation of treatment variables proxying for refugee exposure—both refugees in general, and specific DRC refugees, understood specifically to be Tutsi-Banyamulenges in round 5, but not 3. In order to do so, correct spatial information is necessary for both respondents and refugee settlements. The next sections first outline the geocoding that is available for respondents, and the steps taken to connect respondents to the proximity of refugee settlements.

7.2.3 Geocoding: AidData

The Afrobarometer dataset is geocoded by AidData (BenYishay et al., 2017), seeing that Afrobarometer alone does not provide accurate geospatial data. The geocoding by AidData is thus done post-survey, meaning that there are some limitations on the precision worth discussing. The geocoding follows a double-blind coding system to establish the coordinates of the Afrobarometer Enumeration Areas, meaning that x and y coordinates for the Enumeration Area where the respondent was interviewed is available as a unique variable. However, The methodology is described in more detail in (BenYishay et al., 2017). Figure 9 presents the unique geolocation for rounds 3 and 5 expressed as the number of survey locations geocoded per country.

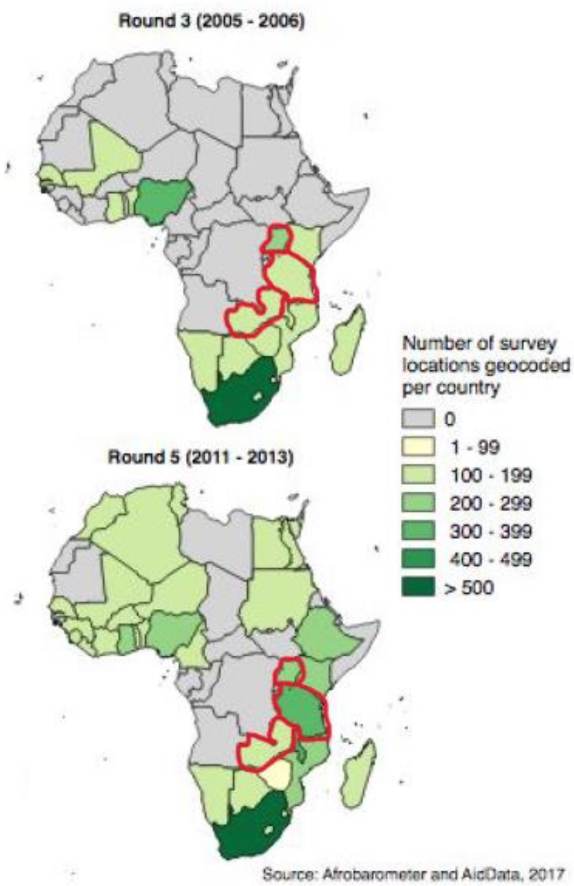

Figure 9. Modified from (BenYishay et al., 2017)

As I want to analyse changes within as well as between countries across the two rounds, the number of geolocations coded have some implications. Recall Dinesen & Sønderskov (2015) who had geodata able to pinpoint Danish respondents with an accuracy of 80 meters. They found that the effect of ethnic diversity on social trust vanished when the distance was between 180-250 meters. If they would compare the results with e.g., German respondents whose geospatial precision was four times less, 320 meters, they would find no effect in Germany. This could give rise to the conclusion that Germany differed from Denmark, although that may not be the case. The geocoded Afrobarometer dataset is nowhere nearly as precise, but the same logic applies. If I were to introduce more granular variables for the countries where it would be possible, entirely different effects might occur and comparison would be rendered useless. Thus, the most imprecise geolocation of respondents at any one round or country effectively establishes the most disaggregated level of geospatial *variables* I can use. The geocoding of Afrobarometer by AidData allows geospatial variables, but no more granular than approximately 100 geographical units per country.

7.2.2 Global Administrative Areas (GADM)

The geocoded Afrobarometer dataset includes variables at different levels such as region, sub-region, district, and counties. However, what constitutes e.g., a district in one country does not necessarily constitute a district in another. District may not even be used as an administrative unit in some countries at all, which is often a problem when working with African data (South et al., 2020). Moreover, the administrative units change over time. E.g., until 2011, the number of districts in Zambia was 72, but have gradually increased to 116 as of 2018. The administrative unit as a variable would thus have changed between rounds 3 and 5, rendering comparison faulty. To overcome this, Global Administrative Areas (GADM) coding is used. This ensures that the administrative units are constant across rounds.

At GADM level 1, Uganda has 56 administrative units; Tanzania has 30; and Zambia has 10. At level 2, Uganda has 166; Tanzania has 186; and Zambia 116. Recall the amount of unique Enumeration Areas geocoded by AidData: the least granular number of Enumeration Areas among rounds and countries is somewhere between 100-199 per country. One could therefore risk respondent's precision of geolocation to be larger than the GADM2 unit and it would consist of less than 30 respondents. It goes without saying that operationalization on such a level is not representative, and I thus base my operationalization on GADM level 1 (henceforth admin1) for all countries, despite there being different number of admin1 units across countries.

Due to the time constraints of the thesis, I utilise the GADM coding in the Ethnic Stratification dataset which builds on rounds 3 and 5 of Afrobarometer, kindly provided by Roland Hodler (Hodler et al., 2020). However, as I am interested in differentiating between respondents that live in proximity of a refugee camp and those who do not, I need to manually code the presence of a refugee camp based on a settlement dataset. It is therefore not sufficient to know the GADM code, but the actual locations the GADM codes represent, which the ES dataset does not include. As GADM codes changes between versions, it is not given which code corresponds to what location. I compare the respondent frequency of ES GADM codes to the respondent frequency of location names in the PRIO Geocoded Afrobarometer dataset and the original Afrobarometer dataset geocoded by AidData. I then use the GADM 4.1 shapefile in QGIS to verify that both location names and GADM codes in fact do correspond, before manually labelling location names on the GADM variables in ES. For a more detailed explanation of the process, sources of uncertainty, and the steps taken to overcome them, see supplementary information in appendix A. The process is depicted in Figure 10.

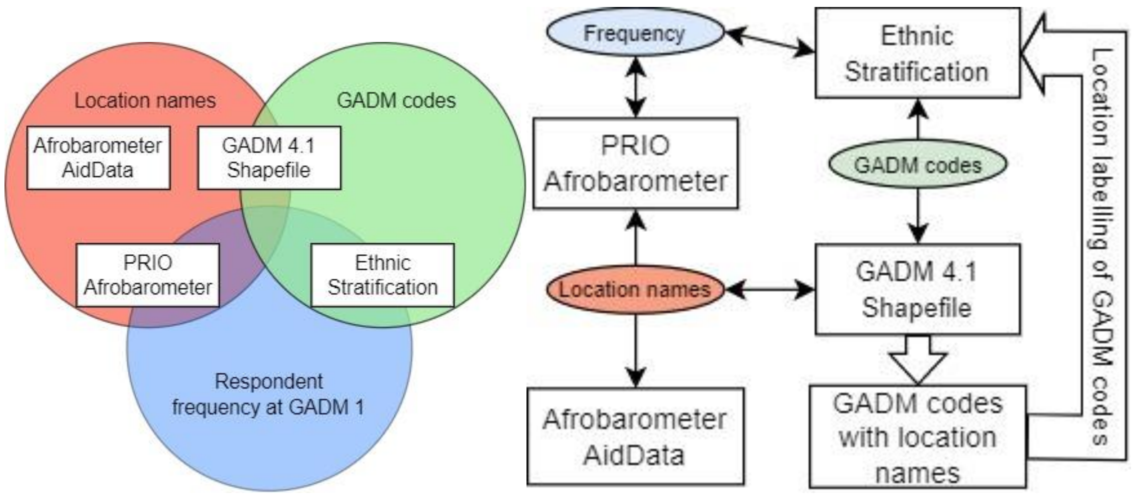

Figure 10. Own model. Data overlap for GADM correction and information flows for coding. Squares=datasets, coloured circles=information.

This process ensures that each respondent is given a standardised and identifiable geolocation which is instrumental for creating treatment variables able to denote a treatment group as well as a control group.

7.3 Operationalisation of spatial treatment variables for refugee exposure

Seeing that I present two hypotheses, the first asks how the exposure of refugees overall affect social trust in the host-community, and the other how ethnic change in the refugee group will influence this, different variables must be created. I create a total of five, the two first are constructed so that one variable can tests the general H1, and the other is a disaggregated version suited for the ethnicity based H2. Both of these are categorical variables based on administrative units, whereas the last three are continuous variables of a computed distance matrix.

7.3.1 Treatment variable 1: Geo-Refugee settlements at admin1

In order to create a good treatment variable capturing refugee exposure, geospatial information on refugee settlements is necessary to test the hypotheses.[4] The Geo-Refugee dataset (Fisk, 2014) provides the solution. Geo-Refugee builds upon UNHCR Location and Demographic Composition data and information from supplemental UNHCR resources, Reliefweb, and news sources. Unlike the UNHCR Settlement Database, Geo-Refugee contains annualised data on settlements. i.e., unique rows for settlement-year, with each settlement having multiple entries in the dataset.

The Geo-Refugee dataset also includes longitude and latitude variables on any given UNHCR refugee settlement from 2000-2019, which allows for plotting the refugee settlements spatially in GIS. Isolating camps only in Uganda, Tanzania, and Zambia between 2005 and 2013 creates a spatial overview of all existing refugee camps in the time frame, as depicted in figure 12. The importance of period accurate data is evident when I overlay these Geo-Refugee settlements with the current UNHCR People of Concern GIS settlement file (UNHCR, n.d.). Despite that both datasets include exclusively UNHCR settlements, the overlaps are surprisingly few.

[4] Spatial information on refugee settlements can be provided directly from the UNHCR. The problem with this data is that it is updated frequently, meaning that correct information between 2005-2013 is unavailable as it does not include annualised information on settlement establishment date or close date, only currently active settlements. This makes the UNHCR spatial data unfit, as temporally sensitive analysis is critical for a difference-in-difference design.

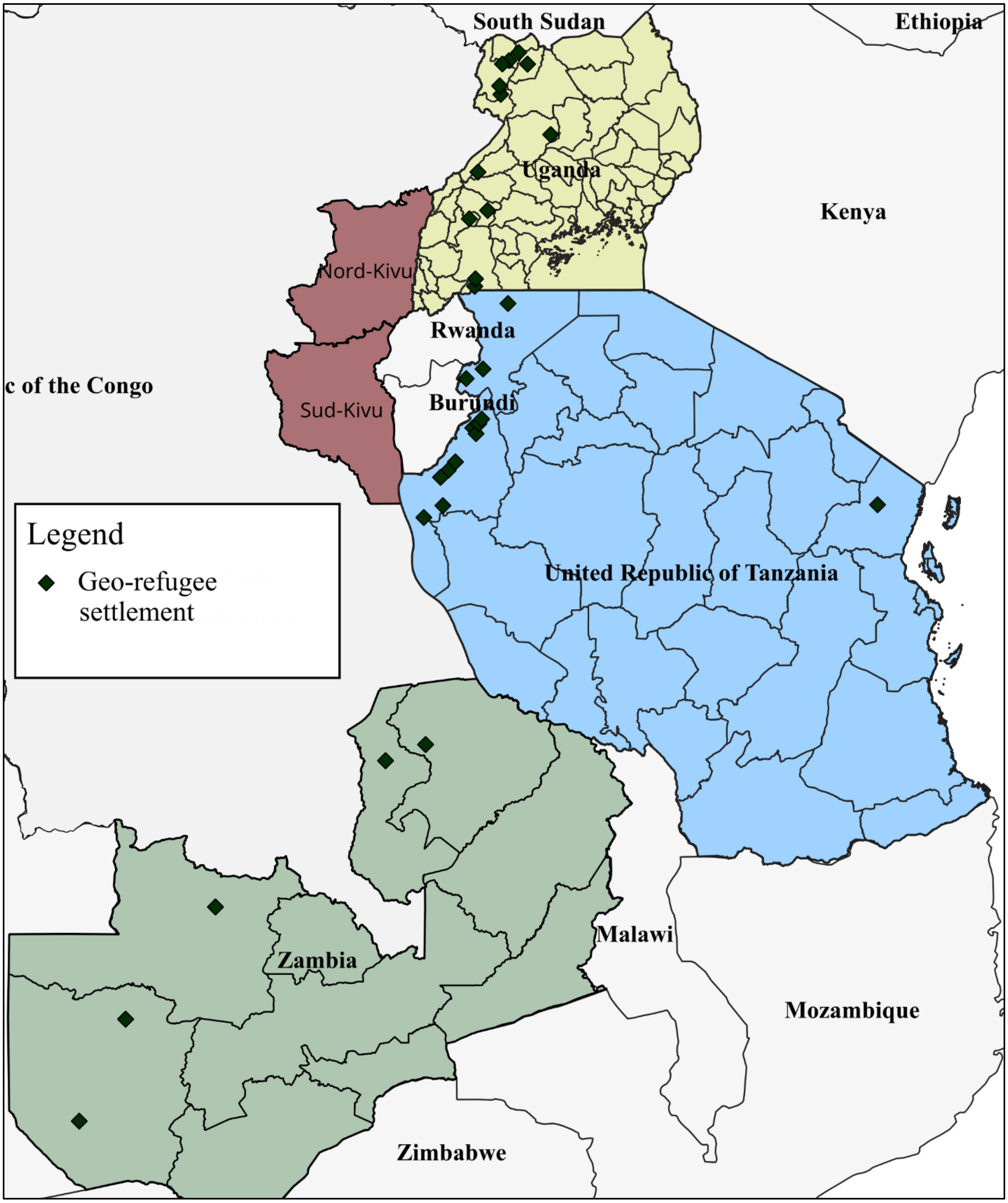

Figure 11. Own visualisation made with QGIS using GADM and Geo-Refugee (Fisk, 2014). Note: Admin1 names and codes removed for clarity.

I construct the dummy variable *Admin1 refugee settlement* by cross referencing my master data GADM codes with GADM1 and GeoRefugee in QGIS. I code every admin1 unit *without* a refugee settlement as 0 and every admin1 unit *with* a refugee settlement as 1. This variable thus functions as a proxy for refugee exposure and can be used to test H1, whether refugees have a negative impact on social trust in the host-society.

7.3.2 Treatment variable for exposure to DRC refugees

The first treatment variable can be viewed as a dummy for exposure to refugees overall, given the precise nature of Geo-Refugee both temporally and spatially. What Geo-Refuge does not do is provide data on the ethnicity of the refugees for the settlements. UNHCR very seldom provides any information on the ethnicity of the refugees in the majority of their settlements. Such data is usually non-existent at the settlement level. For all intents and purposes, refugee settlements are a black box what concerns ethnic composition.

The Ethnicity of Refugees dataset (Rüegger & Bohnet, 2018), as discussed in section 6.2 on the cases, provides data on the majority ethnic refugee group between neighbouring countries with unique rows for each country-dyad, i.e., country of origin and country of arrival, as well as year.

This information is, as stated, used as a base assumption about DRC refugee ethnicity. The majority ethnic group fleeing from DRC to Uganda, Tanzania, and Zambia prior to 2010 are all different ethnic groups. However, from 2010 the ethnic majority group changes to Tutsi-Banyamulenges for all three countries[5]. This is consistent with the development of the Kivu conflict, where the Tutsi rebellion effectively ends in 2009 and is congruous with the UNHCR infographic (UNHCR, 2010b). These three elements combined “fixes” the operationalisation of refugee ethnicity for DRC refugees. This means that with some level of confidence—95% according to Rüegger & Bohnet (2018)—I assume that the majority ethnic group changes for the DRC refugees from 2010.

Identifying which refugee settlements that are primarily populated by Tutsi-Banyamulenges is not as straight forward, nor entirely possible. The Ethnicity of Refugees dataset does not provide geodata. Thus, it is a matter of qualitative assessment. I primarily use the UNHCR Operational Data Portal to access documentation for each single refugee settlement

[5] Note that no data exists for Tanzania prior to 2010

in the Geo-Refugee dataset. The UNHCR documentation usually provide an overview of the total number of refugees and number of refugees by country of origin. I use this UNHCR documentation to assess what percentage the DRC refugees compose in any given camp, or at the minimum if they are the majority group between 2010 and as close to the second measure as possible, that is, 2012-2013.

As the assumption of refugee ethnic group change is already an *assumption* and is prone to a margin of error, I choose to be strict in this assessment. Based on this qualitative assessment, I construct a separate Geo-Refugee Dataset consisting only of the settlements where DRC refugees are explicitly the majority group at the time of documentation. Settlements with significant, but non-majority DRC populations has therefore been excluded. Such settlements include the Rhino Camp in Uganda, and the majority of settlements in Tanzania where the only attainable data is far more recent.

The strict selection of refugee settlements to be included lessens the validity of my operationalised variable of Tutsi-Banyamulenge refugee exposure but increases the variable's reliability. Table 2 includes the table of my qualitative assessment, including a list of each refugee settlement from Geo-Refugee, percentage of DRC refugees, documentation source and year, notes, and whether the camp is ultimately included in the dataset, sorted by country.

Table 2. Qualitative assessment of refugee settlements from GeoRefugee, 2010-2013

Country	Included	Settlement name	Total population	DRC population	DRC %	Majority group	Documentation and year	Note
Uganda	*Included*	Kyaka II	22680	20059	88,44 %	DRC	UNHCR, 2014	From 2012, individual settlement
		Kyangwali	36713	*	*	DRC	UNHCR, 2019a	2017 number, 2019: DRC pupulation at 97%
		Nakivale	60992	30573	50,13 %	DRC	UNHCR, 2014b	Divided into 79 villages across 185km2
		Oruchinga	5212	2250	43,17 %	DRC	UNHCR, 2014c	From 1994, mostly Hutus from Rwanda
		Rwamwanja	52207	52185	99,96 %	DRC	UNHCR, 2014d	From 2012, individual settlement
	Excluded	Rhino Camp	63370	*	<1%	South Sudan	UNHCR, 2016	"62,536 (99%) refugees are from South Sudan"
		Imvepi	57831	18	0,03 %	South Sudan	UNHCR. 2019b	
		Kiryandongo	64362	284	0,44 %	South Sudan	UNHCR, 2022	
		Madi-Okollo	*	*	*	*	None after 2006	Rhino Camp located in Madi-Okollo district
		All exclusively South-Sudan border settlements:						
		Ikafe settlement						
		Nyumazi transit center						
		Pakelle/Adjumani settlement						
		Palorinya						
Zambia	*Included*	Meheba	21905	10341	47,21 %	DRC	UNHCR, 2018b	Local Integration area and refugee area
		Mayukwayukwa	13044	6667	51,11 %	DRC	UNHCR, 2018a	Local Integration area and refugee area
	Excluded	Kala	*	<40 000	*	DRC	UN, 2010	Repatriation, closes in 2010. Others moved to Meheba
		Mwange	*	<40 000	*	DRC	UN, 2010	Repatriation, closes in 2010. Others moved to Meheba
		Nangweshi				Angolan	Shimo, 2006	Closes in 2006
Tanzania	*Included*	Nyarugusu	67400	65000	96,44 %	DRC	UNHCR, 2017c	"Prior to April 2015, the camp hosted 65,000 DRC and 2,400 Burundian Persons of Concern."
	Excluded	Mtabila	35 322	*	*	Burundi	UNHCR, 2013	Closes in 2012 after orderly return to Burundi
		Mtendeli	47296	*	*	Burundi	UNHCR, 2017a	States opening date in 2016
		Nduta	13055 (126740)	*	*	* (Burundi)	UNHCR, 2015; 2017b	States opening date in 2015
		Karago	5500	0	0,00 %	Burundi	Verney & Clark, 2005	Closes in April 2005
		East Burundi border:		East Rwanda Border:		No documentation:		Note
		Lukole				Kanembwa		Located between Nduta and Mtendeli
		Mbuba		Mwisa		Lugufu		
						Mkugwa		Located between Nduta and Mtendeli

Table 2. Qualitative assessment of refugee settlements to be used for the operationalisation of treatment variables on majority ethnic group change. Note that the numbers are based on documentation year.

With this spatial information plotted visually in GIS, presented in fig 12, I construct the second treatment variable *Admin1 settlement*. This variable effectively disaggregates the first treatment variable into three distinct respondent groups[6]. Admin1 units without any settlement are coded as 0, henceforth referred to as the “no settlement group”. Admin1 units with non-DRC refugee majority settlement is coded as 1, henceforth referred to as the “other settlement group”. Lastly, the admin1 units with DRC-majority refugee settlement is coded as 2, henceforth referred to as the “DRC settlement group”. This last group of host-country respondents are of the greatest interest. The variable will thus be the main independent variable for the analysis. This is what this thesis operationalises as the true *treatment* in the quasi-experimental design. Understood in terms of DD, the variable provides the treatment group as well as two distinct control groups. This requires some ingenuity for the specific DD estimates, which is greatly discussed in section 8.2.

[6] Alternatively, a distinct dummy variable could be constructed akin to the H1 treatment variable, where only the DRC-majority settlement group is coded as 1. These two separate dummy variables could, however, not be included in the same model, due to multicollinearity or otherwise influencing coefficients. This is particularly important for logistic regressions where covariates influence coefficients far more than in OLS regression (Mehmetoglu & Jakobsen, 2022, p. 193) They would thus have to be tested with separate models. This approach would be the best suited for answering H1, but less so for hypothesis 2. The reason being that comparing the effects of variables *between* models is prone to be erroneous.

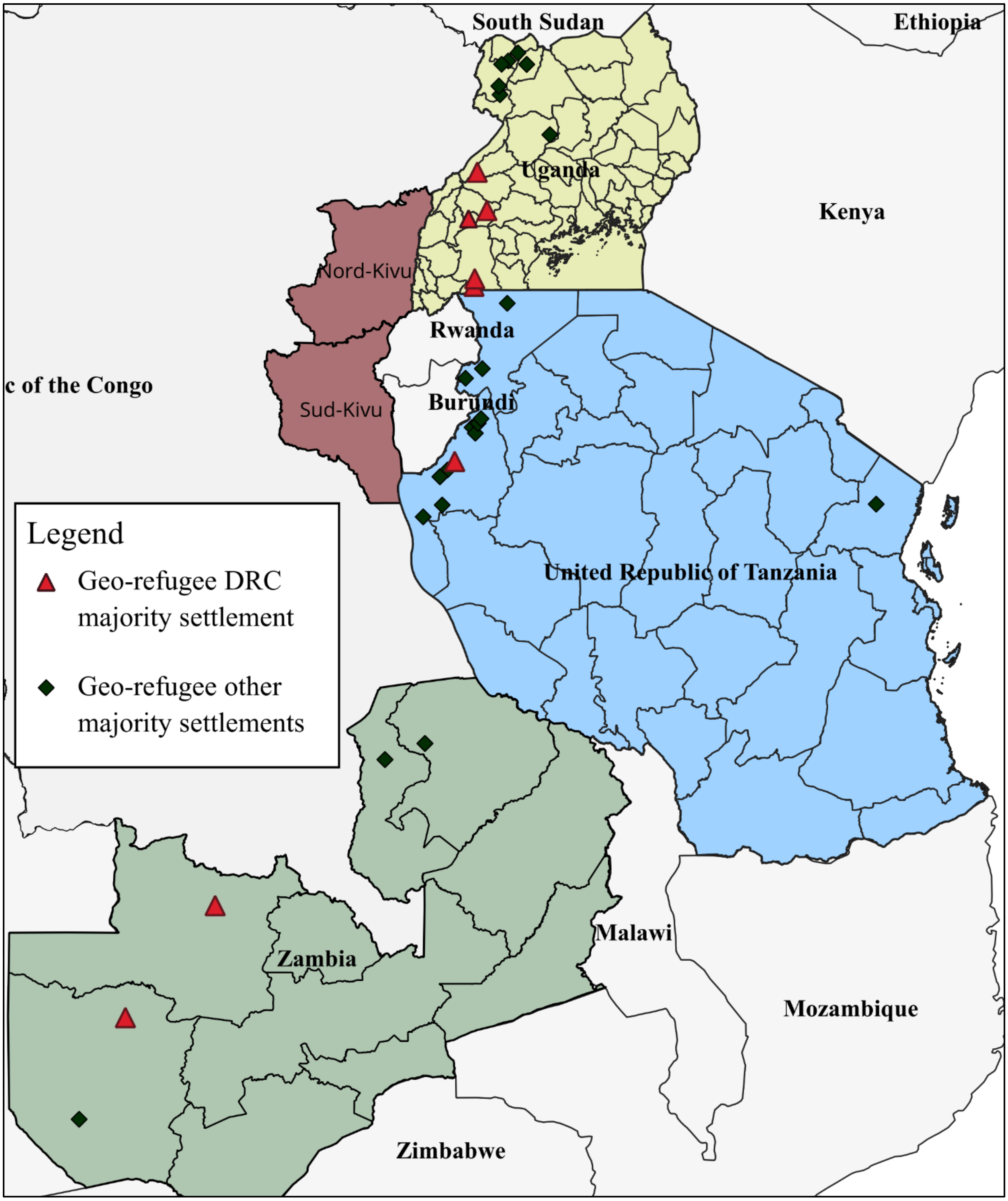

Figure 12. Own visualisation made with QGIS, GADM, and Geo-Refugee (Fisk, 2014).
Note: GADM1 names and codes removed for clarity.
Note: The two DRC majority settlements at the border between Tanzania and Uganda are both located in Uganda.

Analysed as a categorical dummy set, this variable effectively functions as a proxy for DRC Refugee exposure, and in extension Tutsi-Banyamulenges only in round 5 for all three countries. For round 3 however, the ethnicity of refugees for this specific respondent group is not Tutsi-Banyamulenges, but Hutus in Uganda and Tabwas in Zambia (see fig. 6-8 in section 7.2).

7.3.3 Continuous treatment variables

As depicted in figure 11 and 12, the relative size of the admin1 units varies greatly across cases. Uganda—which is the smallest of the three—has 56 admin1 units, where Zambia have only ten. As a result, the likelihood of the admin1 variables accurately reflecting *exposure* to refugees greatly varies between the countries. To address this issue, I introduce two supplementary treatment variables that are continuous in nature.

The geocoded Afrobarometer dataset includes longitude and latitude variables for each respondent, based on the identified Enumeration Area. This allows plotting each respondent in GIS software, Using vector analysis tools, I compute a distance matrix between Afrobarometer respondent and the single closest GeoRefugee settlement. I construct one variable for closest distance between respondent and closest "any" settlement, corresponding to H1; one "other" refugee settlement, whereas the second variable is the distance to the closest DRC settlement. These variables thus correspond to H1 and H2, respectively.

The GIS software does not account for time variables. In order to avoid respondents receiving a distance value to a settlement that is not active yet, or have closed, I separate the two Geo-Refugee subsets again. Both the "any", "other", and the "DRC" subsets of GeoRefugee are split into settlements active in 2005, and settlements active in 2010. Then only round 3 respondents' distance to closest settlement is calculated to both settlement types in 2005, whereas the round 5 respondents' distance to closest settlement is calculated only for the three settlement categories for 2010 only. Appending the round 3 respondents' distance with the round 5 respondents' distance for each category leaves three distinct, spatially, and temporally accurate variables: One for the closest "any" settlement, one for closest DRC settlement, and one for the closest "other" settlement.

These third treatment variables solve some inherent problems with the admin1 based categorical variables: The admin1 based treatment variables are ultimately flawed, as Geo-Refugee settlements often lie at the border of an admin1 unit. This will lead to respondents living rather far from a settlement to be coded as 1, whereas a respondent living much closer to the settlement, yet in another admin1 unit, will be coded as 0. The variables based on the distance matrix solves this problem perfectly due its continuous nature. Yet, as Dinesen & Sønderskov's study (2015) showed, distance may only be relevant up to a certain point. As all respondents are given a distance value and, given the limited N of respondents living in proximity to a settlement as opposed to those who don't, the effects may not be clearly visible. The distance variables are therefore greatly skewed towards the higher values, and I have log-transformed them to reduce the skewness (Skog, 2005, p. 311).

A second problem of the distance matrix is that it is calculated for all respondents as a linear distance to the *closest* refugee settlement. Often, that is a settlement not situated in the respondent's country. Whereas I argued this was one of the shortcomings for the categorical admin1 variables, it is opposite in this case. I expect relations, trade, travel, and interaction to be normal between administrative units of a country, but not across country borders. The type and sizes of the settlements also varies. Some are spatially quite confined or closed off, whereas others span well above 100 square kilometres, such as the Nakivale settlement in Uganda with 79 villages (UNHCR, 2014b). For settlements such as Nakivale, respondent distance to centre of the settlement can be high, yet the respondent may very well live within the confines of the settlement itself.

A third problem of utilising such a continuous distance variable is the precision of the Geocoding (see section 7.1.4, above). The number of unique geolocations differs greatly between rounds and countries. For the first and second categorical treatment variables, which are admin1 based, the geocoding is less granular than the least granular AidData geocoding, and thus the difference in geocoding precision across both countries and rounds becomes irrelevant. This is the strength of the two first treatment variables. The caveat is the sizes of the administrative units differing between countries, meaning that the potential distance between respondent and refugee settlement likewise differs. So then, does the validity of the measure. I therefore opt to construe both types of variables and run parallel analyses with them.

7.3. Control variables

Some socioeconomic control variables are available in both rounds of Afrobarometer. The social control variables of age and gender are included in the models to overcome possible sampling biases, as the data will be analysed unweighted. Previous literature have documented an effect of education on social trust (Güemes & Herreros, 2019). A categorical variable of education is available in Afrobarometer, with 10 categories. I recode them into four distinct categories, denoted by the highest completed category: less than primary schooling (combined “no primary schooling” and “some primary schooling”); primary schooling; secondary schooling; and post-secondary schooling.

Economic variables are more difficult, particularly for round 3. Due to modelling consistency, I only include economic variables available for both rounds. These include one subjective continuous variable of present living conditions, ranging from very bad to very good, for a total of five categories.

Some economic indicators are also filled in by the interviewer in conjunction with a field supervisor (Carter, 2008; Park, 2015). These are dummy variables coded as 1 for the presence of a health clinic, piped water, and electricity grid in the respondent’s Enumeration Area, or Primary Sampling Unit. Statistical correlation tests indicated that an index was viable between piped water and electricity grid (see appendix B for pairwise correlations matrix between all variables). However, the likelihood of one being present but not the other is higher than the added benefit of an index consisting of only two values. Combined with a Cronbach’s Alpha score of 0.68 I choose not to construct the index and rather include both variables individually (Ringdal & Wiborg, 2022, p. 158).

7.4. Descriptive statistics

Table 3 presents descriptive statistics for the entire sample, meaning Uganda, Tanzania, and Zambia combined. As I also analyse country-subsamples, the same descriptive statistics but for each country is included in Appendix B.

Table 3. Descriptive Statistics (N = 10518)

	Freq.	Mean	SD	Min.	Max.
Generalised social trust		.14	.34	.00	1.00
Round (dummy)					
Round 3	4695	.44			
Round 5	5823	.55			
Admin 1 refugee settlement (dummy)					
No settlement	8562	.81			
Any settlement	1956	.19			
Admin1 refugee settlement					
No settlement	8594	.82			
DRC settlement	753	.07			
Other settlement	1171	.11			
Distance matrix any settlements 2005/2010 (logged)		12.14	1.01	5.82	13.94
Distance matrix other refugee settlements 2005/2010 (logged)		12.38	.92	5.82	13.94
Distance matrix DRC refugee settlements 2005/2010 (logged)		12.54	.90	8.58	14.05
Male (dummy)		.50	.50	.00	1.00
Age		35.33	13.26	18.00	99.00
Present living conditions		2.52	1.17	1.00	5.00
Enumeration Area: Health clinic (dummy)		.58	.49	.00	1.00
Enumeration Area: Electrical grid (dummy)		.38	.49	.00	1.00
Enumeration Area: Piped water system (dummy)		.36	.48	.00	1.00
Education categorical					
Less than primary schooling	3055	.29			
Primary schooling	5216	.50			
Secondary schooling	1301	.12			
Post secondary schooling	946	.09			

8. Statistical Methods and Empirical Strategy

This chapter establishes a consistent and logical methodological approach that can be utilised for analysing the cases combined and separately. The method section is organised in two parts, seeing that the methodology must be adapted for each hypothesis. The overarching quantitative research design is framed around difference-in-differences and a unique variety of the triple difference (DDD) estimator to explore ethnic variable moderation. Given the difficult nature of the research design, a stepwise analysis will be employed drawing on multiple quantitative methods, seeing that some methods are better suited for either hypothesis, but not both. A brief reminder of them is in order:

H1: Refugees negatively impact the level of generalised social trust.

H2: The negative effect of refugees on generalised social trust in the host community is moderated by ethnic differences.

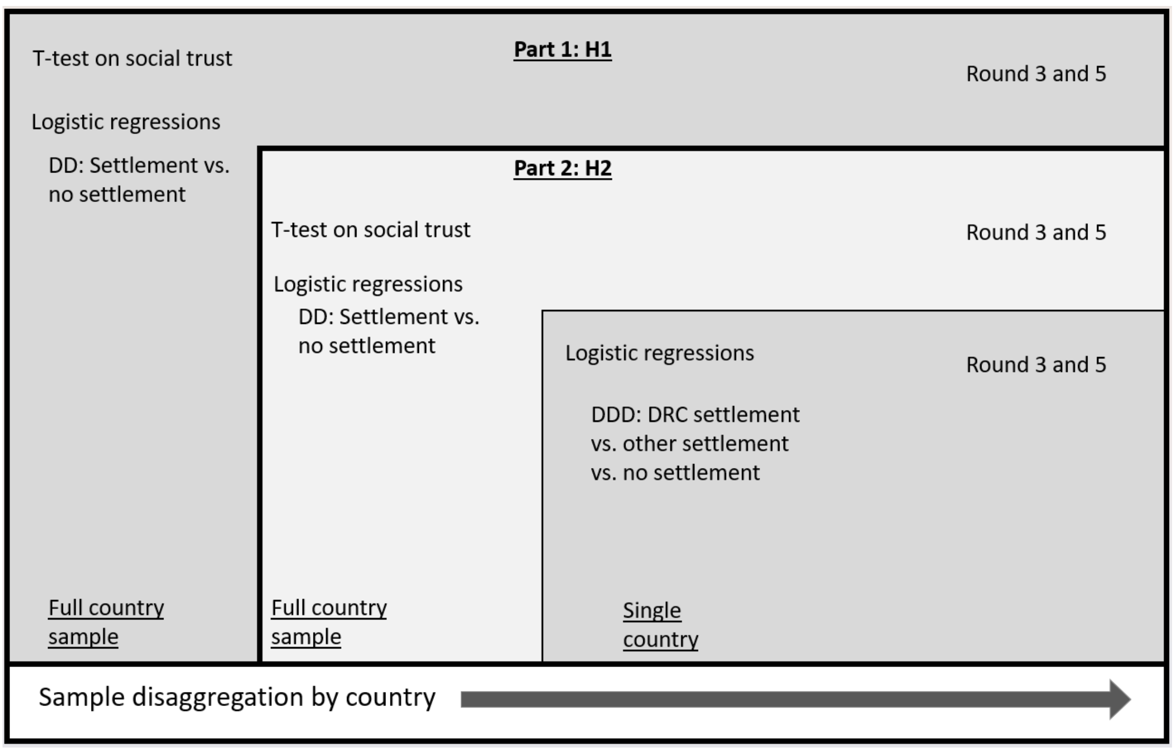

Figure 13. Own model. The stepwise disaggregation of refugee settlements at admin1 and countries.

With the treatment variables I have constructed, H1 is fundamentally an easier hypothesis to test. Keeping ethnicity out of the mix allows for across-case analysis, as the assumption of the hypothesis is that refugees negatively affect social trust regardless of either the ethnicity of the refugees, ethnic composition of the host-community, or the differences between them. Testing H2 is a far more complex and problematic task, as the underlying assumption is that the moderating effect will be different depending on host-society.

Part 1 primarily tests H1 with the fully aggregated sample and the dichotomous admin1 refugee settlement variable. Part 2 will explore the viability of H2. For both parts, the categorical admin1 variables are the primary focus, whereas the continuous distance matrix variables will be used as parallel verification, offering robustness given their complimenting strengths and weaknesses. For part 2, a more disaggregated approach is necessary, as both the ethnic change in the refugee settlement subgroup needs to be accounted for as well as between country effects.

For both part 1 and 2, independent two-sample t-tests will be conducted on trust by round, which is a good starting point for analysing change in subsets of any sample (Skog, 2005, p. 183). This lays the groundwork for logistic regression models. These models are used for part I's difference-in-difference design (DD) separating between admin1 units with refugee settlements, and those without. The specifics of the DD method are discussed in the next section.

For part 2, exploring the viability of ethnic moderation between refugees and social trust, the DD design must be expanded. Part 2 abandons the dichotomous separation of admin1 units with and without settlements, and instead introduces the categorical disaggregation of the admin1 refugee settlement variable, as described above. Recall that DRC-majority settlements in all three countries experience the same change in the dominant refugee group from 2010. This is what this thesis operationalises as H2 treatment. The admin1 units with DRC settlements are therefore the primary interest. DD is ultimately unfit for analysis of three groups. Instead, I employ my own unique difference-in-difference-in-differences (DDD) design, modified from Gruber (1994); Møen & Olden (2020); and Wooldridge (2007).

The stepwise approach offers several benefits. Firstly, it provides serious robustness and sensitivity tests, as the models are tested on various subsets of the sample. Secondly, it allows for a clear understanding of the *statistical* effect that both exposure to all refugees, and exposure to the DRC refugees specifically have on social trust in the cases combined, and then how the effect differs between them.

The downside of the stepwise approach is that the scope of the analysis exponentially increases. To manage this increased scope, the stepwise disaggregation will primarily focus on one singular case (Uganda) with briefer comments on the effects in the other cases. Uganda is chosen due to its large number of admin1 units with and without DRC refugee group majority settlements and best documentation on country of origin for the settlements closest to round 5. The admin1 units of Uganda are also smaller than the other cases, providing increased likelihood of the admin1 coding of settlement presence to reflect actual exposure to refugees for the respondents. The selection of Uganda represents a easier test, meaning that the thesis cannot confirm the theoretical argument, only dismiss it. Given the limited space of previous research empirically and theoretically—and arguably methodologically—choosing the most likely case for the theoretical argument is best suited for early theoretical testing. A least likely case would provide a harder test, which is more appropriate with well-established theoretical causal mechanisms, which this thesis does not operate with. The following section outlines the process, where particularly the difference-in-difference design is discussed in detail.

8.1. T-tests

The initial and simplest statistical testing involves conducting t-tests on the means of the dependent variable, which is generalised social trust, between different rounds and groups. This step can provide early indications if the changes across the rounds are likely due to random variation or underlying patterns in the subgroups of the sample (Skog, 2005, p. 181). An underlying assumption of the t-test is that the variance is equal in the groups, if not, the t-test must be specified to account for unequal variance (Mehmetoglu & Jakobsen, 2022, p. 47). I test this assumption with Levene’s test, which indicates equal variance between the rounds. This initial analysis sets the stage for further statistical testing, having established whether *something* changes between rounds, and in what groups.

8.2. Logistic regressions

The primary statistical analyses for both part 1 and 2 are logistic regressions, performed in Stata. The model fitting follows the same procedures for both parts. Consider first part 1: From these models, the calculated predicted probabilities are used for difference-in-differences estimates, incorporating both the admin1 settlement group (treatment/control) as well as before-and-after (round 3 and round 5) elements in a singular model. Recall that Afrobarometer *is not* panel data, meaning that respondents are unique between rounds. This is the reason that time-series logistic analysis is impossible.

A key assumption of logistic regression is that each observation is independent of each other, and this is partly met by the nature of the merged dataset where the round-prefix is added for each respondent (Mehmetoglu & Jakobsen, 2022, p. 184). One can nonetheless argue that the respondents from the same Enumeration Area are not independent. Therefore, all models are clustered on the standard errors for each Enumeration Area, providing robust standard errors. To estimate the coefficients necessary for difference-in-differences, the model must be specified as to account for both round and group via an interaction term, formalised as:

$$logit(y = 1) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{1i} + \beta_3 X_{1i} X_{2i}$$

Where social trust is represented by y; the intercept baseline log-odds as β_0; the dummy variable for round 5, where round 3=0 and round 5=1, as X_{1i}; and X_{2i} represents the dummy variable for admin1 refugee settlement presence. The interaction term between X_{1i} (round) and X_{2i} (admin 1 settlement) is represented by $\beta_3 X_{1i} X_{2i}$. This can be expressed as:

$$logit(y = 1) = \beta_0 + \beta_1 \text{round } 5_{1i} + \beta_2 \text{admin1 refugee settlement}_{1i} + \beta_3 \text{round } 5_{1i} \text{admin1 refugee settlement}$$

Interpreting interactions in the logit scale is complex and can be prone to multiple misinterpretations. Norton et al (2004) investigate 72 economic journal articles between 1980 and 2000 with nonlinear interaction terms, finding that none of the studies interpreted the coefficient correctly. A common way to overcome the problems of interpreting logit coefficients is to transform the logits to probabilities (Mehmetoglu & Jakobsen, 2022, p. 193). This can be done by calculating margins for the interaction in the model, which provides the predicted probability for all combinations of the interaction term, meaning round (0-1) and admin1 settlement (0-1) (Mehmetoglu & Jakobsen, 2022, p. 133). In other words, the predicated probability of generalised social trust=1 is presented for all four unique combinations of rounds and groups. These predicted probabilities can be expressed as a percentage probability of trust=1 for each variable combination. It must be noted that these are not predictive models, or necessarily inference models, but explorative. Predicted probabilities still represents the most tangible way to interpret and obtain estimates useful for DD.

8.3 Analysis Part I: Difference-in-differences

The predicted probabilities obtained from the calculated margins of the logistic regression serves as difference-in-difference estimates.[7] For part I of the analysis, the control group is the admin1 settlements where no refugee settlements are present, meaning that only the Afrobarometer respondents in these admin1 units constitute the control group (C). The treatment group (A) is the respondents living in an admin1 unit where *any* refugee settlement is present between the points of measure, as outlined in table 4.

Table 4. H1 Difference-in-differences model

		Round 3 (2005/2006)	Round 5 (2012/2013)
H1 Control group	Admin1 with no refugee settlement	C1	C2
H1 treatment group	Admin 1 with refugee settlement	A1	A2
	Difference	A1-C2	A2-C2

The strength of DD as opposed to only analysing changes in the outcome (for this thesis, that would be round 5), is that the DD calculation takes the pre-treatment values of both groups into account and calculating it out, also known as the first difference (Round 3). By subtracting each group's post-measure, the second difference (round 5) from the pre-measure (round 3), I am left with only the change, or rather, the difference. When this is calculated for both groups, I am left with two sets of differences. The final step to obtaining the DD estimator, or Average Treatment Effect (ATE) is subtracting this value for A from C, formalised as:

$$DD = (A2 - A1) - (C2 - C1)$$

[7] Note that all DD estimates in this thesis are calculated manually from the predicted probabilities of the logistic models as explained in the prior section. The user-written Stata command "Diff" can alternatively be used as the analysis tool, directly providing the difference-in-difference estimates (Villa, 2016). These provide the exact same DD estimates as the manually calculated DD estimates from my logistic model's predicted probabilities, both with and without covariates. I choose not to use the Diff command as it only supports DDD with four groups, whereas I have six groups in part 2 of the analysis, as the next section discusses.

For the specific part 1 groups, the equation can be expressed as

$$DD = (Admin1\ with\ refugee\ settlement\ POST - Admin1\ with\ refugee\ settlement\ PRE) - (admin1\ with\ no\ refugee\ settlement\ POST - admin1\ with\ no\ refugee\ settlement\ PRE)$$

I am hesitant towards using the term *treatment group* in this part of the analysis because there isn't necessarily an increase in the refugee population for all cases, and the majority of the settlements existed prior to the first measurement of social trust. The "treatment group" is better understood as refugee settlement baseline effect. The observant reader will likely point out that the equal trends assumption discussed earlier is therefore violated. Such a reading is correct but needs clarification: If H1—refugees negatively impact social trust—is rejected, then the assumption of equal trends holds. If H1 is retained, the assumption of equal trends will most likely have been violated as the effect the model tests would already began prior to the first measurement. It is inherently problematic for an assumption of a model to be dependent on the very outcome it seeks to test.

Arguably, DD is therefore unfit for this part of the analysis. I do, however, perform such an analysis and the reasoning is twofold: As stated in the research design and case selection chapter, I have foremost chosen the natural experiment the refugee population's ethnicity and ethnic change in mind, given the difficult task of finding natural experiments. Put differently, the DD design of part 1 suffers so that the DD estimates for part 2 can be as precise as possible. Why then use DD for part 1 when the assumption of equal trends arguably is violated? Partly due to modelling consistency across the parts to uphold the reliability of the thesis, and partly because the awareness of the problems allows for utilising these DD estimates not as treatment effects, but baseline effects to test H2, where actual "treatment" occurs, as understood in quasi-experimental terminology.

In Part 1 of the analysis, the "treatment" group (A) refers to respondents living in *any* administrative unit with a refugee settlement. These settlements encompass diverse ethnicities and nationalities of refugees (including those from DRC) across the three different host countries, ensuring that the effect observed is not specific to a particular group of refugees and neither dependent on specific characteristics of the host-community. More accurately, the "treatment" represents the baseline condition of having a refugee settlement, as the establishment of these settlements occurred prior to round 3 of data collection. Conceptually, this baseline can be understood as follows: If the presence of a refugee settlement (A1) is associated with lower levels of trust compared to the no-settlement group (C1), one would expect this difference to be amplified in round 5 (A2-C2), indicating a further development of the relationship.

8.4. Analysis Part II: Disaggregated admin1 refugee settlement groups: Tripple difference

Whereas part 1 tests H1, it is first in part 2 that the *moderating* effect of ethnicity is explored. This part continues with the same set of logistic analyses as outlined above, but with the disaggregated categorical refugee settlement variable instead of the dummy: where the no settlement group is coded 0; the DRC refugee settlement group is coded as 1; and the other settlement group as 2.

Consider table 5, where the previous admin1 refugee settlements are split into B (baseline control) and T (treatment). B, being the admin1 "other" refugee settlement group, and T, the admin1 DRC refugee settlement group. The DRC settlement group is where *treatment* understood as *ethnic change* in the refugee population occurs. The terminology and methodology might seem complex, but the reasoning is simple: I argue that the DRC settlement group (T) are, by every metric, more similar to the "other" settlement group (B) than to the no settlement group (C). In other words, group B is a better control group for T.

Table 5. Difference-in-differences-differences (DDD) model

		Round 3 (2005/2006)	Round 5 (2012/2013)
Control group (C)	Admin1 with no refugee settlement	C1	C2
Settlement control group (B)	Admin 1 with other refugee settlement	B1	B2
Treatment group (T)	Admin 1 with DRC refugee settlement	T1	T2
	Difference	T1-B1/B1-C1	T2-B2/B2-C2

In the equal trends assumption it is assumed that if the treatment would not occur, relative development of the groups would continue in tandem (Gertler et al., 2016). Would this be the case for the DRC-settlement group (T) compared to the non-settlement group (C)? This equates to the same problem as discussed in the previous section: Only if H1 is rejected would the equal trends assumption hold. Instead, I argue that if refugees have any effect on social trust—be it either positive or negative—the trajectory of development would already differ between the DRC settlement group (T) and the no-settlement group (C) seeing that the settlements already exist.

Why then, include the no-settlement control group? Without establishing and controlling for the baseline societal development of social trust for the no-settlement group (C), any change between T and B would be a meaningless measure: the B group could potentially mirror the change in overall society or run contrary. If C is dropped, T and B is interpreted in a vacuum. If B is dropped, then whatever DD exist between T and C shows nothing of the *moderating* effect of ethnicity. This leaves a conundrum: which differences to calculate the DD estimate from? Consider figure 14 where the y-axis represents a dependent variable and the x-axis represent measurement time.

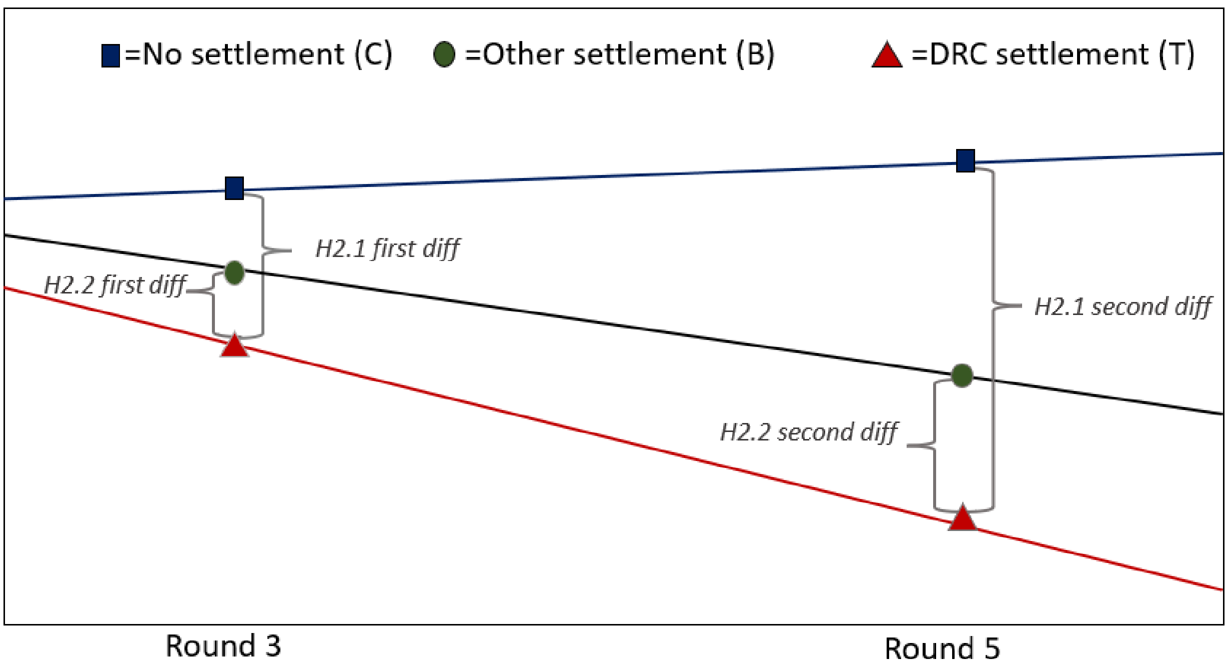

Fig. 14. Theoretical model displaying how the different respondent groups might develop, and the possible differences to calculate. My DDD model incorporates all varieties.

In figure 14, which is a hypothetical model, social trust in admin1 units where there is no refugee settlement has increased. A negative effect occurs between both settlement groups and the dependent variable, but this effect is stronger for T than for B. The various possible differences for T are denoted between the points of measure. If the DD estimate would be based around the H2.1 differences, I would obtain how the specific subset of admin1 refugee settlements differs from overall society. This would indicate nothing what concerns the *relative* and *moderating* effect of ethnic change in the refugee group. Likewise, imagine this plot, but where C is unknown. If I were to use the H2.2 differences, I would ultimately miss the total effect. Figure 15 displays the relationship between the groups spatially (horizontal) and temporal (vertical) as well as where the differences will be drawn.

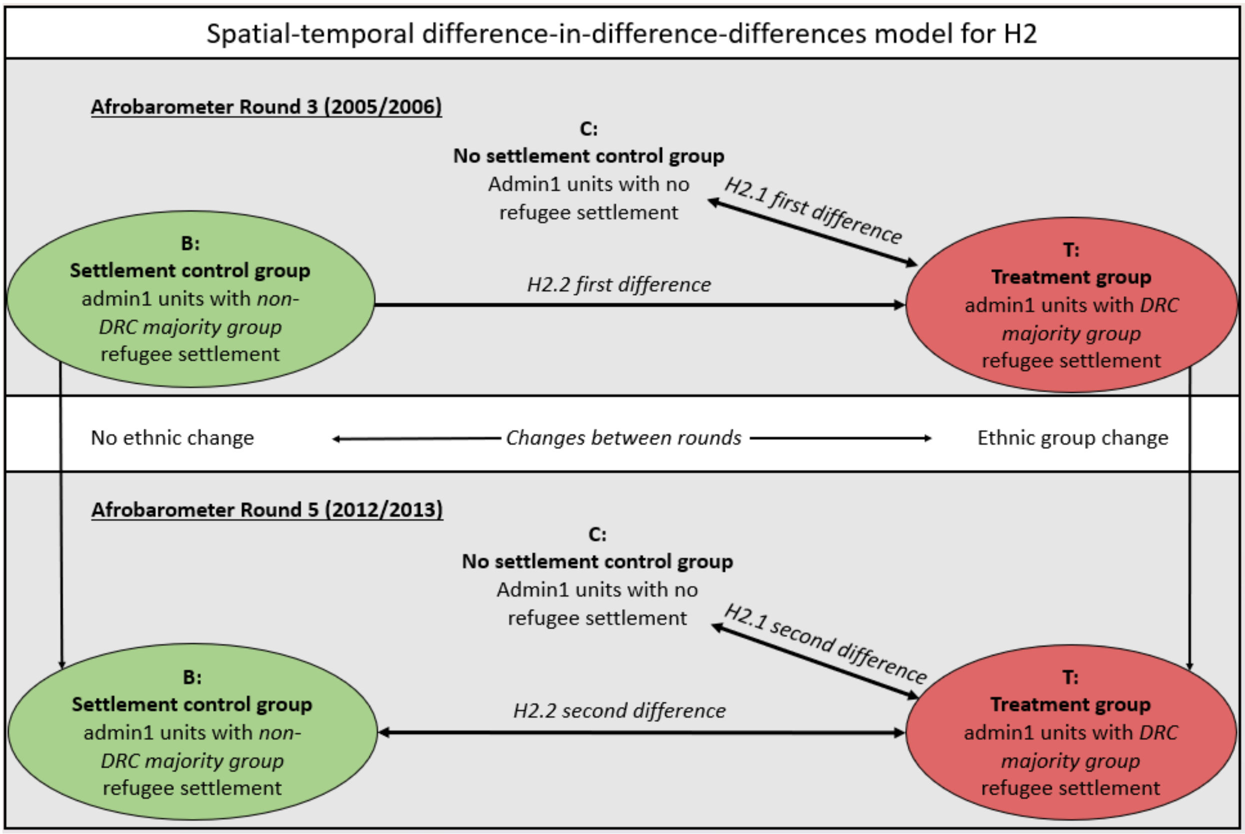

Figure 15. Spatial temporal difference-in-difference-in-difference model

All six measurements and their respective differences are therefore necessary to establish: (i) the effect of refugee exposure on social trust, and (ii) if this effect is different when ethnic changes occur in the refugee group. Only then can one begin to untangle the effect of ethnic change in the refugee group, and how this affects social trust. However, DD is not suited for estimating the effect of treatment with two control groups.

This thesis is not the first instance where this has been a challenge. The solution was introduced by Gruber (1994) which is the difference-in-difference-in-differences (DDD) estimator. Gruber's initial study differed from this thesis, operating with two distinct treatment groups as well as two distinct control groups, surmounting to eight unique measurements with before and after. In reality, this surmounts to two experiments. Which greatly differs from what I am proposing. Møen & Olden (2020) examines the use of DDD in econometrics and the lack of formal treatment received in the literature, pointing to a lecture note by Jeff Wooldridge (2007) as the only somewhat authoritative formal presentation. Møen & Olden explains the DDD estimate as the DD estimation of two existing DD estimates (Møen & Olden, 2020). For my purposes, this must be modified somewhat. Consider their equation:

$$((\overline{y}_{T,B,Post} - \bar{y}_{T,B,Pre}) - (\overline{y}_{C,B,Post} - \bar{y}_{C,B,Pre})) - ((\overline{y}_{T,A,Post} - \bar{y}_{T,A,Pre}) - (\overline{y}_{C,A,Post} - \bar{y}_{C,A,Pre}))$$

Here, two unique treatment groups with two corresponding control groups are analysed. The two different DD estimates are subtracted from each other. I do not have two experiments, but one treatment group (T) and a pseudo treatment-control group (B) as well as keeping the control group (C) in the mix. I argue that the most representative way to capture the *moderating effect* of ethnic change in group T is to first calculate the DD estimate between the other settlement group (B) and the no settlement group (C), precisely as done in the first step to capture the baseline effect of refugees on social trust:

$$DD = (T2 - T1) - (C2 - C1)$$

This ensures methodological continuity as it presents the ATE of the refugee settlement group (B) where no substantial ethnic change in the refugee population occurs, against the no settlement group (C). This can be understood as the disaggregated interpretation of the H1 treatment group for part 2 and is an interesting metric for H1. This DD estimator can then be subtracted from the differences for the DRC settlement group (T), formalised as a modified version of Møen's & Olden's (2020):

$$DDD = (T2 - T1) - ((B2 - B1) - (C2 - C1))$$

This ensures that all six unique predicted probabilities are calculated, and the subsequent ATE captures how the admin1 DRC settlement group differs from the other settlement group *after* the differences of the no settlement group is considered.

Recall the theoretical arguments—which this rather lengthy and overly technical method section may have obscured. The argument presented posits that the ethnic composition of host-society and the ethnic differences between refugee and host society member influences social trust. The models and DDD estimate, however methodically creative, offers no statistical testing of either theoretical claim. These models only build on the descriptive fact that the ethnicity of the DRC refugee group settlements changes between rounds.

8.5. Disaggregated sample by country

Inferring *why* and *how* this ethnic change have an effect requires a more disaggregated approach. This step is therefore to run the same logistic models from the previous section on more disaggregated samples, split by country, and calculating the DDD for each specific country. If the theoretical argument is correct, the DDD estimates ought to be different between the countries. Seeing that the majority ethnic group changes to Tutsi-Banyamulenges in all the countries between the measurements, differing effects across the cases are not due to the specific refugee group, but attributes of the host-society. Performing the logistic regressions with the categorical settlement variable for country specific samples provides unique DDs and DDDs for each country.

9. Results Part I: Refugees' impact on social trust

Summary statistics have already shown that social trust decreases in the overall sample of Uganda, Tanzania, and Zambia. Recall that the wording of the questions differs from round 3 to 5, which could explain the change. A significant t-test would then mean that *something* has affected the social trust between rounds. As round is the time-variable, I use a two-sample t-test specifying social trust by round. First for the entire sample; then for only admin1 units without refugee settlements; and finally, only for admin1 units with any refugee settlement, i.e., both other refugee majority settlements and DRC-majority settlements. The results are presented in table 6.

Table 6. Part I T-tests on social trust

	Round 3		Round 5		Results	
	Mean	*SD*	*Mean*	*SD*	*t*	*p*
Full sample	.141	.348	.134	.340	1081=1.103	.314
Admin1 has no settlement	.128	.005	.132	.004	8813=-.499	.617
Admin1 has refugee settlement	.190	.393	.1464	.353	1982=2.646	0.07

The first two-sample t-test is conducted for the entire three-country sample of 10 815 respondents to compare the mean social trust between round 3 (*M* = .141, *SD* = .348) and round 5 (*M* =.134, *SD* = .340). The t-test showed no significant difference between the rounds, *t*(10813) = 1.103, *p* = .314. This indicates that although the mean negatively changes between the rounds, it may be due to random variation and chance. As I theorise that it foremost is in proximity to refugees that social trust will change, the results are favourable towards the hypothesis testing as it establishes that there is no statistically evident reason to believe that social trust changes systematically in the sample overall.

Recall that the respondents living in admin1 units *with* refugee settlements are relatively few compared to those who don't: 1992 out of the 10815. I perform the same two-sample t-test only for that subsample. The t-test shows different results between round 3 (*M* = .190, *SD* = .393) and round 5 (*M* =.1464, *SD* = .353). The t-test shows a statistically significant and negative difference at 1% level between the rounds, *t*(1982) = 2.646, *p* = .008, and the difference is greater than 0, *p* = .0041, meaning the change is not due to random chance. There is therefore a significant difference between the two measurements of trust for admin1 units with refugee settlements, and this is not likely due to chance.

T-tests are nonetheless arguably not robust and solid estimators to draw conclusions from. The next step of the analysis is thus to perform logistic regressions with the three treatment variables fit for H1, as to better map the relationship between refugees and social trust in the host-community.

9.1. Part I Logistic models and difference-in-differences

For these results, all three countries and both rounds are analysed in a full sample. These full sample models constitute the difference-in-differences design for H1. First, I present two simple models including only the dummy variable for round interacted with a treatment variable: one model per treatment variable. In model (1) the interaction of round (temporal) and the dummy variable for refugee settlement at respondent's admin1 unit (spatial), allows for the calculation of the four discrete predicted probabilities necessary to obtain the DD estimates outlined in the spatial temporal model in fig 15.

In model (2) the logged distance to closest settlement variable is interacted with round. Partly due to the exclusive mutuality of strengths and weaknesses between the categorical and continuous variables, and partly due to the clarity it offers, I present these models in the same table. Model (2) therefore serves as a robustness test, confirming or disconfirming the observed effect in model (1). It is the categorical factor treatment variable in model (1) that in and out of itself offers the difference-in-differences for the cases combined.

Models (3) and (4) are the same model specifications but fitted with socioeconomic control variables. In is these two last models that the predicted probabilities are calculated, and the DD estimates drawn.

Table 7. Part I logistic models. Uganda, Tanzania, and Zambia

VARIABLES	(1) Trust	(2) Trust	(3) Trust	(4) Trust
Round 5 (dummy)	0.032	-1.848**	0.023	-1.519*
	(0.082)	(0.848)	(0.080)	(0.863)
Admin 1 refugee settlement (dummy)	0.430***		0.352***	
	(0.130)		(0.131)	
Round 5 × Admin1 refugee settlement	-0.334*		-0.364**	
	(0.180)		(0.181)	
Logged distance matrix		-0.262***		-0.205***
		(0.053)		(0.054)
Round 5 × logged distance matrix		0.157**		0.127*
		(0.070)		(0.071)
Male (dummy)			0.069	0.064
			(0.050)	(0.050)
Age			0.002	0.003
			(0.002)	(0.002)
Present living conditions			0.044*	0.040
			(0.026)	(0.026)
Enumeration Area: Health clinic			0.151**	0.114
			(0.074)	(0.076)
Enumeration Area: Electrical grid			-0.328***	-0.311***
			(0.089)	(0.090)
Enumeration Area: Piped water			-0.157*	-0.142
			(0.090)	(0.090)
Education.			-0.237***	-0.208***
			(0.069)	(0.069)
Primary schooling			-0.405***	-0.375***
			(0.107)	(0.107)
Secondary schooling			-0.702***	-0.678***
			(0.136)	(0.136)
Post secondary schooling		-0.262***		-0.205***
		(0.053)		(0.054)
Constant		0.157**		0.127*
		(0.070)		(0.071)
	-1.908***	1.280**	-1.826***	0.656
Observations	(0.063)	(0.629)	(0.138)	(0.653)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.10

Table 7 presents several interesting findings. Of greatest interest is the interaction term in models (1) and (3). Here, the combination of living in admin1 unit with a refugee settlement and being a respondent in round 5 is associated with lower social trust, as compared to the respondents living in admin1 units without refugee settlements. For the first model, including only the interaction term without any control variables, this is significant at the 10% level only. Moreover, these results are confirmed by continuous distance variable which has a positive sign, indicating that when the distance from the closest refugee settlement increases, so does the social trust. A more puzzling result is the logit coefficient for the non-interaction term of Admin1 refugee settlement, which indicates respondents living in admin1 units with a refugee settlement, but only those in round 3, seeing that the round variable is a dummy variable where round 5 equals 1. The coefficient is positively signed, indicating that social trust is higher for those living in admin1 units with refugee settlements in round 3 as opposed to those who do not. The interaction displays another story, where the combination of living in an admin1 unit with a refugee settlement in round 5 is associated with lower social trust, which is significant at the 10% level.

The inclusion of the socioeconomic variables in model (3) changes the significance level for the dummy interaction of admin1 settlement and round to the 5% level. This is interesting, and the logit coefficients of the socioeconomic variables are likewise puzzling, both piped water and the presence of electrical grid is negatively associated with social trust, though piped water only at the 10% level. Mind that neither of these variables are interacted with the round variable, meaning that these coefficients reflect respondents for both round 3 and 5 taken together. The dummy set of education paints a similar picture, where more education increasingly negatively impact social trust, congruous with the findings in Güemes & Herreros (2019).

A possible explanation for these negative associations might be refugee settlements themselves, which may explain the change in significance. The reason this could be is that respondents living close to or inside refugee settlements (recall that particularly some settlements in Uganda span hundreds of square kilometres) often benefit from both schooling and infrastructure, as is the case with the Nakivale settlement (UNHCR, 2014b).

Calculating the predicted probabilities from model (2) enables far more tangible interpretation of all four unique combinations of round and admin1 refugee settlement. Figure 16 presents a marginplot of the predicated probabilities, here understood as the four difference estimates.

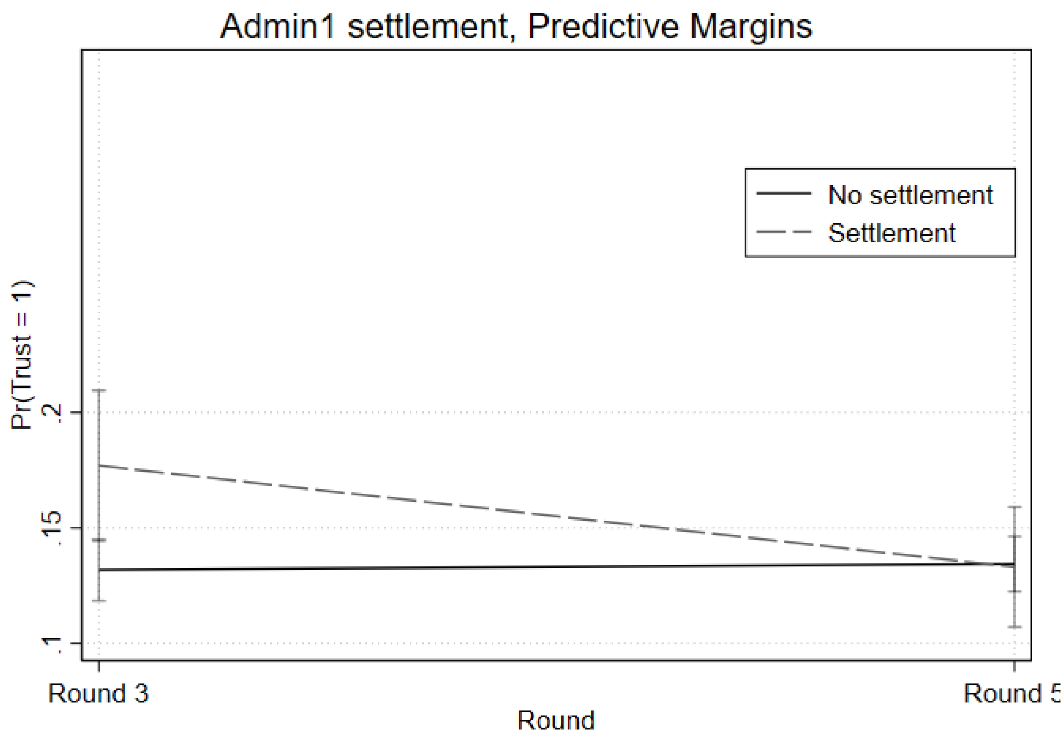

Figure 16. Predicted probabilities for the respondents in admin1 units with refugee settlements and those without.

The plot confirms the interpretation above, but more accurately describes the relationship. Do however note that these are not observations, but predictions based on the model. Interestingly, the predicted social trust is far higher in admin1 units with refugee settlements in round 3, but in the seven years between the measurements this reverts to the point there is no legible difference in predicted probability. Table 8 presents the predicted margins, as well as the DD estimate, Average Treatment Effect.

Table 8. DD predicted probabilities

	Round 3 (2005/2006)	Round 5 (2012/2013)	DD ATE
Admin1 without refugee settlement	.132	.134	-0.046
Admin 1 with refugee settlement	0.18	0.13	

The ATE indicates the relative effect of receiving “treatment”, when both pre and post differences between the two groups are calculated out, indicating that the effect of refugees in the time period results is a decrease of -0.046 predicted probability of trust being 1, or 4.6 percentage points lower.

The Round 3 estimate does not constitute a before measure, but rather a baseline measure where the difference between the groups is expected to compound by round 5. Whatever may have caused the higher probability of social trust in round 3 where refugee settlements existed, have not compounded, but rather reverted. This gives rise to the neither a confirmation nor rejection of H1, but rather stands as an empirical puzzle.

10. Results Part II

The results in part 1 indicates a fairly strong negative relationship between respondents in admin1 units with refugee settlements in round 5 and generalised social trust. It did, however, cause an empirical puzzle: The predicted probability of trust in round 3 was far higher among the respondents in the refugee settlement group than for the respondents in the no settlement group. In round 5, the predicted probability was nearly indistinguishable. Hypothesising outside the scope of the data is a path perhaps best left untread, yet it is evident that refugee presence is associated with higher social trust in round 3. For round 5, the opposite seems evident.

Part II of the analysis seeks to nuance this picture by disaggregating the admin1 refugee settlement group into two distinct groups: One group of respondents living in admin1 units with *DRC majority settlements* that experiences a change in the majority ethnic group to Tutsi-Banyamulenges in all three cases, and the *other settlement* group where no such change is evident. Keep in mind that these settlements are not comprised of DRC refugee majorities, but a wider variety of ethnicities and nationalities. Due to the time constraints the thesis operates in, the ethnic composition in these settlements have not been ascertained. This category of the variable is operationalised as the logical continuation of H1.

10.1 Part II T-tests

Recall that the full sample and no settlement groups in part I had no statistically significant change in the means, however, the respondents in admin1 units with refugee settlements had. Table 9 presents the results of the t-tests.

Table 9. T-test on social trust

	Round 3		Round 5		Results	
	Mean	*SD*	*Mean*	*SD*	*t*	*p*
Admin1 has no settlement	.128	.005	.132	.004	8813=-.499	.617
Admin1 has other refugee settlement	.148	.014	.181	.015	1188=1.564	.119
Admin1 has DRC refugee settlement	.251	.025	.090	.014	792=6.135	.000

I conduct t-tests separately for three groups, as defined by the disaggregated categorical refugee settlement variable. The no settlement group is precisely the same as in part I but included for clarity and ease of comparison. For the other settlement group, the mean of social trust changes from .14 to .18, indicating a surprising and substantial increase. However, the t-test is not significant (p = .119). Although it is close to being significant at the 10% level, it does not meet the commonly preferred 5% level. In contrast, the results from the DRC-majority refugee settlement group show a dramatic negative development of social trust from .14 to .09, which is statistically significant (p = .00). It is therefore probable that this development is what was responsible for the changes observed in the second t-test in part 1, the group with any refugee settlement.

10.2. Part II Logistic full sample models and modified difference-in-difference-in-differences

The models presented in this section mimics the models in part I: it is the full sample consisting of all three cases, Uganda, Tanzania, and Zambia. The model outline follows the same organization, with some changes: the first model (1) is the model fitted with the categorical admin1 settlement variable. However, the *settlement group* is now split between *admin1 DRC settlement* and *other settlement*, whereas the reference category remains the *no settlement* group. This model captures the difference between the respondents living in admin1 units with DRC settlements where a change in the ethnic majority group occurred between round 3 and 5, and the respondents living in admin1 settlements where no ethnic change occurred.

The second model (2) is fitted with the logged distance value to *other* closest settlement. The third model (3) is fitted with the logged distance value to the closest *DRC* settlement. These two models constitute robustness tests for model (1), given that they are perfectly complimentary to the categorical variable in terms of strengths and weaknesses. Models (4) (5) and (6) repeats this model outline, but with added socioeconomic control variables. It is these three final models that the predicted probabilities are calculated from, used as the DD and DDD estimates. The marginplots presented are the outputs of these predicted probabilities.

Table 10. Part II logistic models. Uganda, Tanzania, and Zambia

VARIABLES	(1) Trust	(2) Trust	(3) Trust	(4) Trust	(5) Trust	(6) Trust
Round 5 (dummy)	0.028	2.393***	-6.01***	0.015	2.819***	-5.725***
	(0.082)	(0.907)	(0.92)	(0.080)	(0.896)	(0.946)
Admin1 refugee settlement						
DRC settlement	0.720***			0.655***		
	(0.183)			(0.187)		
Other settlement	0.154			0.062		
	(0.151)			(0.149)		
Round 5 × Admin1 DRC settlement	-1.240***			-1.310***		
	(0.280)			(0.287)		
Round 5 × Admin1 other settlement	0.189			0.192		
	(0.205)			(0.204)		
Distance matrix other settlement (logged)		-0.068			-0.012	
		(0.049)			(0.049)	
Round 5 × Distance matrix other settlement		-0.192***			-0.230***	
		(0.074)			(0.073)	
Distance matrix DRC settlement (logged)			-0.37***			-0.317***
			(0.06)			(0.057)
Round 5 × Distance matrix DRC settlement			0.48***			0.456***
			(0.07)			(0.076)
Male (dummy)				0.067	0.062	0.064
				(0.050)	(0.050)	(0.050)
Age				0.003	0.003	0.003
				(0.002)	(0.002)	(0.002)
Present living conditions				0.046*	0.042	0.044*
				(0.026)	(0.026)	(0.026)
Enumeration Area: Health clinic				0.183**	0.141*	0.152**
				(0.073)	(0.074)	(0.073)
Enumeration Area: Electrical grid				-0.354***	-0.355***	-0.299***
				(0.090)	(0.090)	(0.090)
Enumeration Area: Piped water				-0.143	-0.122	-0.186**
				(0.091)	(0.091)	(0.091)
Education.						
Primary schooling				-0.236***	-0.204***	-0.215***
				(0.068)	(0.069)	(0.069)
Secondary schooling				-0.387***	-0.366***	-0.369***
				(0.108)	(0.108)	(0.108)
Post secondary schooling				-0.702***	-0.691***	-0.677***
				(0.136)	(0.137)	(0.135)
Constant	-1.901***	-0.988*	2.70***	-1.850***	-1.641***	2.124***
	(0.062)	(0.580)	(0.69)	(0.139)	(0.592)	(0.709)
Observations	10,518	10,518	10,518	10,518	10,518	10,518

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.10

The results in table 10 provides a series of clarifications on the evident effects observed in part I of the results. Disaggregating the refugee settlement group nuances the significant and strong results in a variety of ways. It is evident in these models that the effects from part I largely can be explained by the DRC refugee settlement group alone. This is also the only part of the interactions in models (1) and (4) that are significant, and at the 1% level. For both model (2), (3), (5), and (6), the interactions of the logged continuous distance variables are significant at the 1% level. Note that higher values on the distance matrix equals further distance from a refugee settlement, i.e., if they are to confirm the categorical models, their signs ought to be opposite. The signs of the interactions of the distance values in model (2) (3) (5) and (6) all confirm the results of models (1) and (4), providing serious robustness due to the variables' complimentary nature.

Turning to models (4) again, keep in mind that the categorical variable is specified such that it is the no settlement group that is the reference category. The significance levels therefore indicate if the other categories significantly differ from the no settlement group[8]. The models indicate that the *other settlement* group does not. This gives rise to rejecting the null hypothesis, thereby retaining the notion that the social trust in the other settlement group does not systematically differ from the no settlement group. I argued that this subset of refugee settlements is the logical continuation of H1—that refugees negatively impact social trust in the host community. These results alone highlight the importance of ethnic change in the refugee group. Figure 17 presents the predicted probabilities for the six unique combinations of groups and round for model (4).

[8] Specifying the models so that the DRC refugee settlement group is the reference category returns significant coefficients for both other groups and the round 3 variable. This is not surprising, seeing that the DRC settlement group differs so starkly from both other groups for round 3 and 5. The models are specified with the no settlement group as the reference seeing that this is the largest group, over eight times larger than either settlement group. For the predicted probabilities, the specification of the reference group is irrelevant as it predicts all six combinations regardless, returning the same estimates.

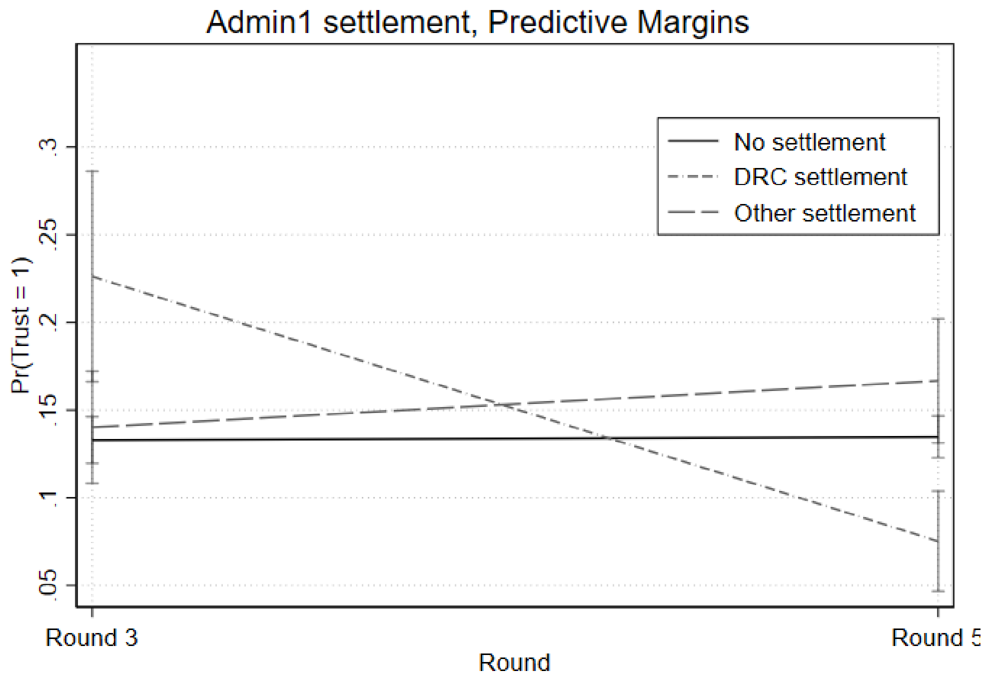

Figure 17. Predicted probabilities for generalised social trust.

Seeing that this is an early study into the moderating effect of ethnic change in the relationship between refugees and social trust in the host community, allow me to foray outside the narrow hallways of statistical significance for a moment. In figure 17, the predicted probabilities of model (4) are plotted. Social trust among respondents in the no settlement group is shown to be rather stable between the rounds. Both the refugee settlement groups are predicted to have higher levels of social trust in round 3. The "other" settlement group is predicted to have developed even higher likelihood of social trust in round 5, running contrary to H1—this model indicates that refugees are associated with a higher level of social trust.

Particularly interesting is however that the DRC settlement group is predicted to have nearly twice the level of social trust than the no settlement group for round 3, which is statistically significant. However, contrary to the other settlement group, no positive change occurs, but rather a dramatic decrease in predicted social trust, which again, is significant. To construe an even clearer picture of the effect of the treatment—understood as refugee group ethnic change in the respondent group living in admin1 units with a DRC refugee settlement—the modified variety of the DDD ATE can be calculated, taking the relative development of all three groups into account.Table 11 presents the DD ATE for the other settlement group and the no settlement group, whereas the DDD estimate is the DRC settlement group where the former DD estimate is subtracted.

Table 11: DDD predicted probabilities

	Round 3	Round 5	DD(D) ATE
No settlement group	.1329	.1347	
Other settlement group	.1402	.1667	.0247
DRC settlement group	.2262	.0752	-.1758

The ATEs indicate that respondents in the other settlement admin1 unit experience an increase of social trust by 2.47 percentage points between the round. The relative effect of living in an admin1 unit with DRC settlements is severely different. When both pre and post measure differences are calculated out along with both other groups' differences, the effect of the change of ethnic refugee groups corresponds to a decrease in 17.5 percentage points for the host community respondents' social trust. This is a substantial amount.

These results definitely add a new dimension of understanding from part I, but the empirical puzzle is far from solved. Evidently something has strongly and positively influenced the social trust levels for these admin1 units prior to round 3 whereas something has caused this to change, not only to revert back to the levels of society overall, but far lower. This thesis, with its framing around ethnic change, lends itself to the belief that this is precisely the reason. This is also the very reason that such interpretation should be taken with extreme caution. An important next step is therefore to analyse each country isolated to map how social trust develops differently for the three respondent groups.

10.3 Part II Logistic models disaggregated by country

The models and DDD estimates presented in this section follow the exact same outline as those in the prior section. The difference being that the models are performed on country-subsamples. Due to parsimony, the most likely case Uganda is presented. The tables with logistic regressions for Tanzania and Zambia are available in appendix D. Table 10 presents the logistic models. Model 4—both for Uganda as well as Tanzania and Zambia—provides the logit coefficients used to calculate the predicted probabilities and the following DDD estimates.

Table 12. Part II logistic models. Uganda

VARIABLES	(1) Trust	(2) Trust	(3) Trust	(4) Trust	(5) Trust	(6) Trust
Round 5 (dummy)	0.001	7.436***	-11.666***	-0.036	8.933***	-10.947***
	(0.107)	(1.519)	(1.614)	(0.103)	(1.460)	(1.609)
Admin1 refugee settlement						
DRC settlement	0.874***			0.785***		
	(0.215)			(0.223)		
Other settlement	0.087			-0.043		
	(0.203)			(0.196)		
Round 5 × Admin1 DRC settlement	-1.007***			-1.034***		
	(0.334)			(0.355)		
Round 5 × Admin1 other settlement	0.328			0.469		
	(0.307)			(0.303)		
Distance matrix other settlement (logged)		0.132*			0.236***	
		(0.079)			(0.075)	
Round 5 × Distance matrix other settlement		-0.629***			-0.758***	
		(0.128)			(0.123)	
Distance matrix DRC settlement (logged)			-0.478***			-0.430***
			(0.078)			(0.074)
Round 5 × Distance matrix DRC settlement			0.970***			0.909***
			(0.134)			(0.134)
Male (dummy)				0.060	0.059	0.061
				(0.073)	(0.073)	(0.073)
Age				0.005*	0.006*	0.006*
				(0.003)	(0.003)	(0.003)
Present living conditions				0.004	-0.005	0.021
				(0.033)	(0.034)	(0.033)
Enumeration Area: Health clinic				0.177	0.249**	0.055
				(0.125)	(0.119)	(0.120)
Enumeration Area: Electrical grid				-0.284**	-0.341**	-0.282**
				(0.144)	(0.147)	(0.142)
Enumeration Area: Piped water				-0.188	-0.179	-0.134
				(0.161)	(0.165)	(0.163)
Education.						
Primary schooling				-0.160*	-0.160	-0.134
				(0.097)	(0.097)	(0.097)
Secondary schooling				-0.067	-0.043	-0.088
				(0.142)	(0.142)	(0.143)
Post secondary schooling				-0.742***	-0.762***	-0.735***
				(0.185)	(0.186)	(0.184)
Constant	-1.692***	-3.155***	4.066***	-1.723***	-4.456***	3.454***
	(0.077)	(0.921)	(0.924)	(0.190)	(0.896)	(0.879)
Observations	4,735	4,735	4,735	4,735	4,735	4,735

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.10

The results in table 10 illustrate a similar picture as the models for the cases combined. The interaction indicates significantly lower trust for respondents in admin1 units with DRC refugees as opposed to those with no settlement, whereas the effect is positive but not significant for the other settlement group. Like the combined case models, the robustness testing provided by the logged continuous distance matrix values confirms the relationship. The distance to DRC settlement interaction behaves as in the prior step: significant at the 1% level and positively signed, confirming the results of models (1) and (4). They are also significant for *other* settlement in model (5), opposite of the DRC settlement distance. Unlike the combined models, they also indicate that in round 3, trust increases with distance to closest settlement.

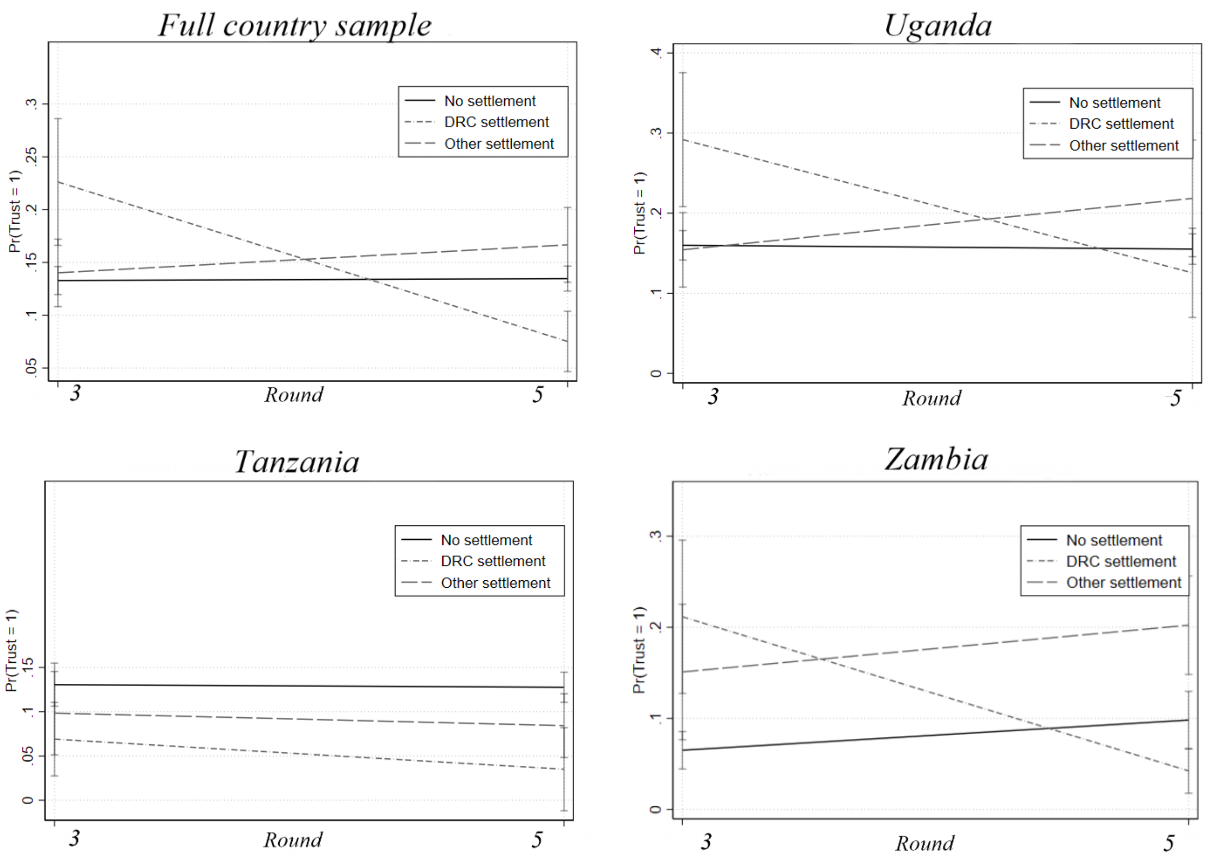

Figure 18. Predicted probabilities for all part 2 logistic models with control variables. All predicted probabilities obtained from country-subsamples models (4). The tables for Tanzania and Zambia can be found in Appendix D.
Note: All graphs specified to identical y axis.

Performing the logistic models for each country individually provides a powerful analytical insight: For both Uganda and Zambia, the majority ethnic refugee group prior to 2010 was different. In Zambia, it was Tabwas for a series of years, whereas it was Hutus in Uganda. For all three countries, the majority ethnic group of DRC refugees are Tutsi-Banyamulenges in 2010. This means that in round 5, conducted in 2012/2013, the DRC majority groups are the same for all countries. An extension of this is that the differences between the effects of this specific DRC group ought not to be that group in and out of itself, but differences in the host community.

Note first how Uganda differs from the full country sample. In round 3, the predicted probability for trust in admin1 units with a DRC settlement is far higher. Though the negative trend is very similar, at round 5 the prediction is barely lower than for the no settlement group. For Tanzania, very few of the coefficients return significant values in any of the models. Neither distance variable is significant but indicates congruous signs. For model 4, which these predicted probabilities are calculated, only the DRC settlement group at round 3, but not round 5 as specified in the interaction, are significant at the 5% level. This is however sufficient to state that Tanzania greatly differs from Uganda at the first point of measurement. Whereas admin1 DRC settlement respondents in Uganda at round 3 had close to .3 predicted probability at round 5, the respondents in Tanzania living in admin1 units with DRC settlements had barely .07. Moreover, despite the interaction term not being significant, the predicted values only draw a small downward trend into round 5.

In Zambia, however, both the DRC settlement group and the other settlement group are significant at the 1% level for round 3. For the interaction term, the same is true for the respondents living in admin1 units with DRC settlements, but not the other settlement group. In fact, this interaction term changes sign from positive in the non-control model to positive when controls are included. However, the round 5 variable, which due to the interaction term indicates those living in admin1 units without refugee settlements at all, is significant at the 1% level. Recall however that Zambia only has ten admin1 units, meaning that whether this variable proxies for refugee exposure is far less likely than either Tanzania or Uganda. All distance matrix robustness tests are significant at the 1% level and signed as to support the categorical variable.

Viewed in tandem, and with some disregard to significance levels, the visual plots tell somewhat similar stories, but not without its caveats. Particularly interesting is therefore that for two of the cases, respondents living in admin1 units with DRC settlements have a significantly higher predicted probability for social trust in round 3, and that this is wholly reverted in round

5. I argue that the likelihood of this being due to the change in the ethnic refugee group is high given the combined results from all the cases. I further argue that the different effects are due to different characteristics in the host community. If these differences can be extrapolated to be ethnicity from these analyses, is more questionable. But, the rather dramatic negative change observed in all countries, save Tanzania that has quite unreliable coefficients, leads me to state that this is highly plausible.

Table 14. DDD ATE for step of the analysis and each country

	DD (other-no)	DDD
Part I full sample	-.0465	
Part II full sample	.0689	-.1758
Uganda	.0689	-.2353
Tanzania	-.0113	-.0227
Zambia	.0183	-.1876

Table 14 presents all the DD and DDD estimates from the analyses, calculated as described in the prior section. As all estimates are predicted probabilities from categorical interaction terms, not every estimate is statistically significant at the 5% level, meaning that the DDD must be viewed with some caution. Apart from Tanzania where no results were statistically significant, both Uganda and Zambia illustrate a stark picture: the change in the refugee group to Tutsi-Banyamulenges—when all differences are calculated out, meaning only the relative difference remains—are associated with severe and substantial decrease in social trust. 23 percentage points decrease, and 18 percentage points decrease for Uganda and Zambia, respectively. The results are not intended for causal claims, but to explore the viability of ethnicity being a moderating factor in the relationship between refugees and social trust—to which there remains no doubt.

10.4 Robustness tests

Due to the nature of my stepwise testing and consequent inclusion of the continuous distance variable models, the effect has withstood rigorous sensitivity and robustness tests already. However, changing the dependent variable to fear of crime can offer further robustness. Although different, the concept captures a sense of trust in society and is highly dependent on social trust (Walklate, 1998). I argued that Newton's (2007) definition of trust was composed of a positive and a negative element. Fear of crime relates to the later. I conducted the same models with OLS regression, seeing that fear of crime is a continuous variable with five categories. The relationship was quite similar inn all of the countries, yet no statistically significant values of the interaction terms was obtained. This is puzzling, though one explanation is that it is foremost the positive element of social trust that the respondents give greatest weight. I do not view these results as inherently problematic, seeing that fear of crime and whether one trusts people arguably is conceptually quite different. The results of the robustness tests can be provided upon request.

11. Discussion

The analysis has a series of implications worthwhile of discussing. The first three of which are methodological. Analysing the complex relationship of ethnicity, refugees, and host-community attitudes of trust in Sub-Saharan Africa is severely difficult due to data availability. Such research depends on precise spatial and temporal data, which limits the already scarce basis. Ethnic fractionalisation indices are inherently problematic and Afrobarometer ethnic data is at this stage not suited to address the topic adequately at a meaningful level of disaggregation. This thesis was written precisely as an endeavour to overcome this, and I argue partly succeeds by drawing on a multitude of datasets, qualitative assessments, methodical ingenuity, and some leaps of faith what concerns assumptions. The use of a framing conflict overcomes a series of problems concerning case selection and helps "fix" central empirical challenges when working with refugee questions in Sub-Saharan Africa, such as ensuring time of arrival is after the first point of measurement. Refugee population statistics from UNHCR or Ethnicity of Refugees (Rüegger & Bohnet, 2018) alone cannot suffice, as repatriations schemes are common and will obfuscate the relative increase in refugees. The use of a framing conflict or event can be employed for other, diverse research questions when researching data-poor cases where temporal precision is vital.

The results obtained with the stepwise disaggregation also shows the importance of analysing disaggregated data, which the empirical research community long since have taken to heart. The insights of the analyses in part I and part II describe the same data, yet the conclusions one would draw from them alone are near polar opposite. This begets the question of what other empirical surprises one would be in for if further disaggregation was performed either spatially or to the refugee settlement host respondent groups. This presents both interesting avenues for further research, but also tells a cautionary tale towards large N across-country analyses on complex research questions.

The second methodological contribution is the opportunities provided by my unique DDD estimator, able to encapsulate two distinct control groups for one singular treatment group. This lessens the burden of the equal trends assumption, enabling analysis of relative change for more unorthodox or seemingly problematic or outright impossible quasi-experiments. To the extent of my knowledge, only Wooldridge (2007) have presented a similar DDD estimator, but as pointed out in Møen & Olden (2020) Wooldridge's equation cannot be utilised as is and is conceptualised for eight groups rather than six. Future research could benefit from expanding the scope of what can constitute quasi-experiments with a similar approach as in this thesis, particularly in less explored and less documented areas such as Sub-Saharan Africa.

The fourth implication is towards the theoretical debate on generalised social trust generation. The analysis shows that generalised social trust in the majority of the sample, which is the no settlement group, is largely stable across the seven years between round 3 and round 5. This lends some support to the psychological perspective of how trust is generated, as argued by Uslaner (2002, 2008). This does come with a central caveat: In two of the three countries analysed, social trust is higher in admin1 units with any settlement, DRC majority or not. This means that my results indicate an association between refugee settlement and higher social trust in the host community. Though few of the "other settlement groups" return statistically significant coefficients. The robust and substantial decrease in social trust for the DRC settlement group where ethnic refugee group change occurs, does however indicate that social trust *can* change. This in turn lends some support to the experiential, rational-choice argument of generalised social trust generation. I initially endorsed both arguments, arguing that substantial changes in one's social surroundings could warrant changes in social trust attitudes, despite such attitudes being stable and predisposed. The analysis may not singlehandedly prove such a reading of the theory, but it provides some evidence that social trust is both stable and prone to change. The extent to which this is confirmed here is nonetheless limited, as Afrobarometer is not panel data.

The most substantial insight of the analysis is arguably what the thesis set out to test: if change in the majority ethnic group would negatively affect social trust in the host community. The results indicate precisely that. It is nonetheless with some caution that claim is made. No variables on either ethnicity or ethnic heterogeneity were introduced for the respondents. The ethnic data available in Afrobarometer did not allow for *directly* testing in- and outgroup theory, and the Ethnic Stratification index was only available for 26 African villages, of which too few were situated in admin1 units with refugee settlements, leading to perfect predictions and consequent omitted results. This is the area which future research is the most needed. The use of the framing conflict and the observed change in the refugee group to a group with different ethnicity somewhat bypasses this. I argued that when such a change occurs, any notions of in- or outgroup views of the refugees based on ethnicity must likewise change. As such, this analysis has operationalised precisely this, and the finding suggest that change of the ethnic group of refugees have negative impacts on trust.

As a piloting study on the moderating effect of changes in ethnic refugee group, the results are none other than striking. There is a robust and negative effect on social trust in the host community where such change occurs, and they are robust between several countries. This gives rise to accepting H2, that the effect of refugees on social trust is moderated by ethnic group change. This is an important early finding in untangling the attitudinal effects for refugee hosting communities and needs to be explored further.

In answering the question: Does the European insights hold? The answer must be: Partly, and if one were to focus only on the *negative* effect of the DRC refugee settlements from round 3 to 5, then yes. This would nonetheless be a case of not seeing the wood for the trees, for the results also present an empirical puzzle: there is an association between refugee settlement and higher social trust in the host community as opposed to where no settlements exist. This is true for the DRC settlements in round 3 as well as the “other” settlements, although the evidence is weaker. For three of the four cases, this settlement group also has a positive trend into round 5, though only statistically significant in Zambia. This insight is entirely made possible by disaggregating the settlements where ethnic group changes occurs. This gives rise to the rejection of H1: refugees *may* in fact have a positive effect on social trust in the host community. When highlighting this aspect of the relationship between social trust and refugees, then no, the European insights do not hold, they run contrary.

Unfortunately, no data is available to assess this positive trend between social trust and "other" settlements further back in time. There are, however, some possible explanations. Several studies (e.g. Kreibaum, 2016; Maystadt & Duranton, 2014; Taylor et al., 2016) have highlighted the positive economic effects refugees and aid can bring to the host-community, which the few economic variables I include as controls not fully can account for. E.g., in the Nakivale settlement, the host population benefits both from education and infrastructure (UNHCR, 2014b). In the greater social trust literature, increased living standards and welfare is strongly associated with high social trust, expressed as the North-South divide. Such an economic explanation of the possible benefits of hosting refugees may therefore shed some light into *why* social trust is higher in refugee hosting admin1 units. Such an explanation can only go so far. Another explanation is precisely the one Dinesen (2012) presented: ethnic heterogeneity. Of course, no such claim is tested in this thesis, and the argument runs opposite to Dinesen's: Due to the greater ethnic diversity and the fact that state borders in Sub-Saharan Africa do not correspond well to ethnic borders, refugees may not represent an outgroup. As Rüegger & Bohnet (2018) find, refugees show a preference to flee where they have ethnic kin. The arrival and presence of refugees, then, can constitute an increase in ingroup members for large parts of the host-community and in extension create more ethnically homogenous communities where social trust generates more easily. This is an exciting possibility, though far outside what this analysis have shown. Untangling precisely such ethnic relationships is the important next step for nascent research on the effects of hosting refugees on attitudes of social trust.

12. Conclusion

Social trust is seen as a vital component for economic, democratic, and social development. Overcoming the low-trust equilibriums currently present in Africa is problematic, and research on the phenomenon is still limited. This thesis has contributed to this research gap by specifically investigating the relationship between hosting refugees and generalised social trust. By theorising a moderating effect of ethnic refugee group change, several fascinating and thus far unexplored empirical insights were made possible. Employing a unique difference-in-difference-in-differences design, I have accounted for various spatially and temporally respondent groups in relation to proximity to specific refugee settlements. Utilising a framing conflict to overcome central empirical challenges have enabled me to analyse the effect of change of ethnic refugee group, indicating substantial, statistically significant, and robust decreases in social trust for the host community members. The results also indicated an empirical puzzle, as the “other” refugee settlements—where no such ethnic change occurs—are associated with high social trust for its host-community. This opens for exciting new research better suited at precisely exploring the mechanisms of why. The early findings nonetheless indicate that the European insights may not hold in an African context.

I introduced this thesis by a cautionary tale of the possible adverse effects of the currently halted UK-Rwanda deal, highlighting that the refugee situation in Sub-Saharan Africa is far more pressing than in the Global North. Such schemes, suggesting essentially outsourcing *all* illegal immigrant could compound the situation. This thesis has furthered the knowledge of the possible negative effects of hosting refugees, particularly when ethnic changes are introduced. Now, refugees hosted in African communities are largely from neighbouring countries, separated by borders that entails both ethnic, historical, and cultural ties. The UK-Rwandan scheme shows no regard for ethnic and cultural bonds. This thesis supports the notion that other European powers, eager to find a solution to their political refugee problems, would do well to consider the detrimental effects of such schemes.

Literature

Afrobarometer. (n.d.). *Afrobarometer Sampling*. Retrieved 6 May 2023, from https://www.afrobarometer.org/surveys-and-methods/sampling/

Afrobarometer. (2005). *Afrobarometer Round 3*. https://www.afrobarometer.org/survey-resource/merged-round-3-data-18-countries-2005/

Afrobarometer. (2015). *Afrobarometer Round 5*. https://www.afrobarometer.org/survey-resource/merged-round-5-data-34-countries-2015/

Alesina, A., & Ferrara, E. L. (2000). Participation In Heterogeneous Communities. *The Quarterly Journal of Economics*, *115*(3), 847–904.

Alix-Garcia, J., & Saah, D. (2010). The Effect of Refugee Inflows on Host Communities: Evidence from Tanzania. *The World Bank Economic Review*, *24*(1), 148–170. https://doi.org/10.1093/wber/lhp014

BenYishay, A., Rotberg, R., Wells, J., Lv, Z., Goodman, S., Kovacevic, L., & Runfola, D. (2017). Geocoding Afrobarometer Rounds 1-6: Methodology & Data Quality. *AidData at William & Mary*.

Bertinelli, L., Comertpay, R., & Maystadt, J.-F. (2022). *Refugees, Diversity and Conflict in Sub-Saharan Africa*. The World Bank. https://doi.org/10.1596/1813-9450-10052

Bjørnskov, C. (2012). How Does Social Trust Affect Economic Growth? *Southern Economic Journal*, *78*(4), 1346–1368. https://doi.org/10.4284/0038-4038-78.4.1346

Callamard, A. (1994). Refugees and Local Hosts: A Study of the Trading Interactions between Mozambican Refugees and Malawian Villagers in the District of Mwanza [*]. *Journal of Refugee Studies*, *7*(1), 39–62. https://doi.org/10.1093/jrs/7.1.39

Carter, D. (2008). *Data Codebook for Round 3 Afrobarometer Surveys in 18 African Countries*. Michigan State University. https://www.afrobarometer.org/survey-resource/merged-round-3-codebook-18-countries-2005/

Chambers, R. (1986). Hidden Losers? The Impact of Rural Refugees and Refugee Programs on Poorer Hosts. *International Migration Review*, *20*(2), 245. https://doi.org/10.2307/2546034

Chandra, K. (2006). What is Ethnic Identity and Does it Matter ? *Annual Review of Political Science*, *9*(1), 397–424. https://doi.org/10.1146/annurev.polisci.9.062404.170715

Collier, P., Honohan, P., & Moene, K. O. (2001). Implications of Ethnic Diversity. *Economic Policy*, *16*(32), 127+129-166.

Davies, S., Pettersson, T., & Öberg, M. (2022). Organized violence 1989–2021 and drone warfare. *Journal of Peace Research*, *59*(4), 593–610. https://doi.org/10.1177/00223433221108428

Delhey, J., & Newton, K. (2005). Predicting Cross-National Levels of Social Trust: Global Pattern or Nordic Exceptionalism? *European Sociological Review*, *21*(4), 311–327. https://doi.org/10.1093/esr/jci022

Denny, E. K., & Walter, B. F. (2014). Ethnicity and civil war. *Journal of Peace Research*, *51*(2), 199–212. https://doi.org/10.1177/0022343313512853

Dinesen, P. T. (2012). Does Generalized (Dis)Trust Travel? Examining the Impact of Cultural Heritage and Destination-Country Environment on Trust of Immigrants: Does Generalized (Dis)Trust Travel? *Political Psychology*, *33*(4), 495–511. https://doi.org/10.1111/j.1467-9221.2012.00886.x

Dinesen, P. T., & Sønderskov, K. M. (2015). Ethnic Diversity and Social Trust: Evidence from the Micro-Context. *American Sociological Review*, *80*(3), 550–573.

'Dozens killed' in DR Congo raids. (2009, May 13). *BBC*. https://web.archive.org/web/20090516021146/http://news.bbc.co.uk/1/hi/world/africa/8049105.stm

Easterly, W. (2007). *The white man's burden: Why the West's efforts to aid the rest have done so much ill and so little good* (1. paperback ed). Oxford Univ. Press.

Ellingsen, T. (2011). Etnicity Matters, but what Kind? A Review of the Quantitative Literature Of Ethnicity And Conflict. In T. G. Jakobsen (Ed.), *War: An introduction to theories and research on collective violence* (pp. 85–104). Nova Science Publishers.

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review*, *97*(01), 75–90. https://doi.org/10.1017/S0003055403000534

Finseraas, H., & Jakobsson, N. (2012). Trust and Ethnic Fractionalization: The Importance of Religion as a Cross-Cutting Dimension. *Kyklos*, *65*(3), 327–339.

Fisk, K. C. (2014). *Geo-Refugee: A Refugee Location Dataset* [Data set]. Harvard Dataverse. https://doi.org/10.7910/DVN/25952

Fisk, K. C. (2019). Camp settlement and communal conflict in sub-Saharan Africa. *Journal of Peace Research*, *56*(1), 58–72. https://doi.org/10.1177/0022343318814588

Freitag, M., & Traunmüller, R. (2009). Spheres of trust: An empirical analysis of the foundations of particularised and generalised trust. *European Journal of Political Research*, *48*(6), 782–803. https://doi.org/10.1111/j.1475-6765.2009.00849.x

Fukuyama, F. (2015). *Political Order and Political Decay*. Profile Books.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice* (Second Edition). International Bank for Reconstruction and Development/The World Bank.

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research*, *39*(5), 615–637. https://doi.org/10.1177/0022343302039005007

Green, A. (2012, October 23). Congolese Refugees Search for New Homes in Uganda. *Voice of America*. https://www.voanews.com/a/congolese-refugees-search-for-new-homes-in-uganda/1531554.html

Gruber, J. (1994). The Incidence of Mandated Maternity Benefits. *The American Economic Review*, *84*(3), 622–641.

Güemes, C., & Herreros, F. (2019). Education and trust: A tale of three continents. *International Political Science Review*, *40*(5), 676–693. https://doi.org/10.1177/0192512118779184

Habyarimana, J., Humphreys, M., Posner, D. N., & Weinstein, J. M. (2007). Why Does Ethnic Diversity Undermine Public Goods Provision? *American Political Science Review*, *101*(4), 709–725. https://doi.org/10.1017/S0003055407070499

Herreros, F., & Criado, H. (2009). Social Trust, Social Capital and Perceptions of Immigration. *Political Studies*, *57*(2), 337–355. https://doi.org/10.1111/j.1467-9248.2008.00738.x

Hodler, R., Srisuma, S., Vesperoni, A., & Zurlinden, N. (2020). Measuring ethnic stratification and its effect on trust in Africa. *Journal of Development Economics*, *146*, 102475. https://doi.org/10.1016/j.jdeveco.2020.102475

Huening, L.-C. (2013). Making use of the past: The Rwandophone question and the 'Balkanisation of the Congo'. *Review of African Political Economy*, *40*(135), 13–31. https://doi.org/10.1080/03056244.2012.761603

Hutchison, M. L., & Johnson, K. (2011). Capacity to trust? Institutional capacity, conflict, and political trust in Africa, 2000–2005. *Journal of Peace Research*, *48*(6), 737–752. https://doi.org/10.1177/0022343311417981

Jacobsen, K. (2002). Can refugees benefit the state? Refugee resources and African statebuilding. *The Journal of Modern African Studies*, *40*(4), 577–596. https://doi.org/10.1017/S0022278X02004081

Jeannet, A.-M. (2020). Immigration and political distrust in Europe: A comparative longitudinal study. *European Societies*, *22*(2), 211–230. https://doi.org/10.1080/14616696.2019.1694162

Knack, S., & Keefer, P. (1997). Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *The Quarterly Journal of Economics*, *112*(4), 1251–1288. https://doi.org/10.1162/003355300555475

Kreibaum, M. (2016). Their Suffering, Our Burden? How Congolese Refugees Affect the Ugandan Population. *World Development*, *78*, 262–287. https://doi.org/10.1016/j.worlddev.2015.10.019

Lewis, J. D., & Weigert, A. (1985). Trust as a Social Reality. *Social Forces, 63*(4).

Logan, C., & Torsu, A. K. (2022). Ties that bind? Evidence of unity and division in 34 African countries. *Afrobarometer Dispatch Papers, 516*, 1–23.

Lundåsen, S. W. (2023). Engagement in Civil Society and Different Forms of Social Trust in the Aftermath of the European Refugee Crisis. *Nonprofit and Voluntary Sector Quarterly*, *52*(1), 243–259. https://doi.org/10.1177/08997640221077143

Maystadt, J.-F., & Duranton, G. (2014). *The Development Push of Refugees*. 43.

Maystadt, J.-F., & Verwimp, P. (2014). Winners and Losers among a Refugee-Hosting Population. *Economic Development and Cultural Change*, *62*(4), 769–809. https://doi.org/10.1086/676458

McLaren, L. (2012). Immigration and Trust in Politics in Britain. *British Journal of Political Science, 42*(1), 163–185. https://doi.org/10.1017/S0007123411000251

McLaren, L. (2017). Immigration, national identity and political trust in European democracies. *Journal of Ethnic and Migration Studies*, *43*(3), 379–399. https://doi.org/10.1080/1369183X.2016.1197772

Meer, T. van der, & Tolsma, J. (2014). Ethnic Diversity and Its Effects on Social Cohesion. *Annual Review of Sociology*, *40*(1), 459–478. https://doi.org/10.1146/annurev-soc-071913-043309

Mehmetoglu, M., & Jakobsen, T. G. (2022). *Applied statistics using Stata: A guide for the social sciences* (Second edition). SAGE.

Mishler, W., & Rose, R. (2001). What Are the Origins of Political Trust?: Testing Institutional and Cultural Theories in Post-communist Societies. *Comparative Political Studies*, *34*(1), 30–62. https://doi.org/10.1177/0010414001034001002

Møen, J., & Olden, A. (2020). *The Triple Difference Estimator* [[Discussion paper]]. Norwegian School of Economics. Department of Business and management Science. https://openaccess.nhh.no/nhh-xmlui/bitstream/handle/11250/2652016/0120.pdf?sequence=1

Moscona, J., Nunn, N., & Robinson, J. A. (2017). Keeping It in the Family: Lineage Organization and the Scope of Trust in Sub-Saharan Africa. *American Economic Review*, *107*(5), 565–571. https://doi.org/10.1257/aer.p20171088

Newton, K. (2007). Social and Political Trust. In R. J. Dalton & H.-D. Klingemann (Eds.), *The Oxford Handbook of Political Behaviour* (pp. 342–361). Oxford University Press.

Newton, K., & Zmerli, S. (2011). Three forms of trust and their association. *European Political Science Review*, *3*(2), 169–200. https://doi.org/10.1017/S1755773910000330

Norton, E. C., Wang, H., & Ai, C. (2004). Computing Interaction Effects and Standard Errors in Logit and Probit Models. *The Stata Journal: Promoting Communications on Statistics and Stata*, *4*(2), 154–167. https://doi.org/10.1177/1536867X0400400206

Obermeier, A. M., & Rustad, S. A. (2023). Conflict Trends: A Global Overview,. *PRIO Paper*.

Ochieng, Z. (2017). *Congolese Tutsi in the Troubled DRC equation*. Reliefweb. https://reliefweb.int/report/democratic-republic-congo/congolese-tutsi-troubled-drc-equation

Park, C. (2015). *Data codebook for a Round 5 Afrobarometer survey in 34 African countries*. Michigan State University. https://www.afrobarometer.org/survey-resource/merged-round-5-codebook-34-countries-2011-2013-last-update-july-2015/

Posner, D. N. (2004). Measuring Ethnic Fractionalization in Africa. *American Journal of Political Science*, *48*(4), 849–863. https://doi.org/10.1111/j.0092-5853.2004.00105.x

Putnam, R. D. (2007). E Pluribus Unum: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies*, *30*(2), 137–174. https://doi.org/10.1111/j.1467-9477.2007.00176.x

Ringdal, K., & Wiborg, Ø. (2022). *Lær deg Stata* (2nd ed.). Vigmostad & Bjørke.

Ritzer, G., & Stepnisky, J. (Eds.). (2014). *Sociological theory* (Ninth edition, international student edition). McGraw-Hill Education.

Rothstein, B., & Uslaner, E. M. (2005). All for All: Equality, Corruption, and Social Trust. *World Politics*, *58*(1), 41–72. https://doi.org/10.1353/wp.2006.0022

Rüegger, S. (2019). Refugees, ethnic power relations, and civil conflict in the country of asylum. *Journal of Peace Research*, *56*(1), 42–57. https://doi.org/10.1177/0022343318812935

Rüegger, S. (2021). *Ethnicity of Refugees Dataset Codebook* (Version 2021) [Data set]. http://www.icr.ethz.ch/data/er.

Rüegger, S., & Bohnet, H. (2018). The Ethnicity of Refugees (ER): A new dataset for understanding flight patterns. *Conflict Management and Peace Science*, *35*(1), 65–88. https://doi.org/10.1177/0738894215611865

Shimo, K. (2006, November 20). Zambia's Nangweshi Camp closes after last group of Angolans move out. *UNHCR*. https://www.unhcr.org/news/zambias-nangweshi-camp-closes-after-last-group-angolans-move-out

Skog, O.-J. (2005). *Å forklare sosiale fenomener: En regresjonsbasert tilnærming* (2. utg., rev. og utv., 2. oppl). Gyldendal norsk forlag.

Sønderskov, K. M., & Dinesen, P. T. (2016). Trusting the State, Trusting Each Other? The Effect of Institutional Trust on Social Trust. *Political Behavior*, *38*(1), 179–202. https://doi.org/10.1007/s11109-015-9322-8

South, A., van der Walt, A., & Dicko, A. (2020). *Administrative boundaries—Making other data more operationally usefull*. afrimapr. https://afrimapr.org/blog/2020/admin-boundaries/

Stearns, J. (2012). *North Kivu: The background to conflict in North Kivu province of eastern Congo*. Rift Valley Institute.

Stolle, D., Soroka, S., & Johnston, R. (2008). When Does Diversity Erode Trust? Neighborhood Diversity, Interpersonal Trust and the Mediating Effect of Social Interactions. *Political Studies*, *56*(1), 57–75. https://doi.org/10.1111/j.1467-9248.2007.00717.x

Taylor, J. E., Filipski, M. J., Alloush, M., Gupta, A., Rojas Valdes, R. I., & Gonzalez-Estrada, E. (2016). Economic impact of refugees. *Proceedings of the National Academy of Sciences*, *113*(27), 7449–7453. https://doi.org/10.1073/pnas.1604566113

Turton, D. (1997). Introduction: War and Ethnicity. In D. Turton (Ed.), *War and Ethnicity* (pp. 1–46).

UN. (1951). *Convention relating to the Status of Refugees*. United Nations Conference of Plenipotentiaries on the Status of Refugees and Stateless Persons convened under General Assembly resolution 429 (V) of 14 December 1950.

UN. (2010, October 29). UN shuts down camps in Zambia after Congolese refugees go home. *UN News*. https://news.un.org/en/story/2010/10/357552

UNHCR. (n.d.). *UNHCR People of Concern* [UNHCR Core GIS Data]. Retrieved 12 April 2023, from https://data.unhcr.org/en/geoservices/

UNHCR. (2010a). *Convention and Protocol Relating to the Status of Refugees*. UNHCR.

UNHCR. (2010b). *REFUGEES, RETURNEES AND IDPs DEMOCRATIC REPUBLIC OF CONGO* [Fact sheet]. UNHCR. https://data2.unhcr.org/en/documents/details/81938

UNHCR. (2013). *UNHCR Tanzania—Situation Report Mtabila Camp closure November 26—December 02, 2012*. UNHCR. https://data2.unhcr.org/en/documents/details/48435

UNHCR. (2014a). *Kyaka II Fact Sheet 2014* [Fact sheet]. UNHCR. https://data2.unhcr.org/en/documents/details/48486

UNHCR. (2014b). *Nakivale Fact Sheet 2014* [Fact sheet]. UNHCR. https://data2.unhcr.org/en/documents/details/48488

UNHCR. (2014c). *Oruchinga Fact Sheet 2014* [Fact sheet]. UNHCR. https://data2.unhcr.org/en/documents/details/48491

UNHCR. (2014d). *Rwamwanja Fact Sheet 2014* [Fact sheet]. UNHCR. https://data2.unhcr.org/en/documents/details/48493

UNHCR. (2015). *Nduta camp profile* [Camp Profile]. UNHCR. https://data2.unhcr.org/en/documents/details/34169

UNHCR. (2016). *Uganda Rhino Settelement, Arua—UNHCR OPERATIONAL UPDATE* [Situation report]. UNHCR. https://data2.unhcr.org/en/documents/details/52165

UNHCR. (2017a). *TANZANIA - Mtendeli Camp Profile (31 October 2017)* [Camp Profile]. UNHCR. https://data2.unhcr.org/en/documents/details/60872

UNHCR. (2017b). *TANZANIA - Nduta Refugee Camp Situation—3RP KEY INDICATORS MONITORING (31 July 2017)* [Situation report]. UNHCR. https://data2.unhcr.org/en/documents/details/58835

UNHCR. (2017c). *TANZANIA - Nyarugusu Camp Profile—31 March 2017* [Camp Profile]. UNHCR. https://data2.unhcr.org/en/documents/details/56082

UNHCR. (2018a). *Mayukwayuka Refugee Settlement* [Fact sheet]. https://data2.unhcr.org/en/documents/details/67306

UNHCR. (2018b). *UNHCR Zambia: DRC Situation—Meheba settlement Profile as of 30 November 2018* [Settlement Profile]. UNHCR. https://data2.unhcr.org/en/documents/details/67305

UNHCR. (2019a). *Uganda—Refugee Statistics April 2019—Kyangwali* [Settlement Profile]. UNHCR. https://data2.unhcr.org/en/documents/details/70042

UNHCR. (2019b). *Uganda—Refugee Statistics May 2019—Imvepi* [Settlement Profile]. UNHCR. https://data2.unhcr.org/en/documents/details/70038

UNHCR. (2019c). *Uganda—Refugee Statistics September 2022—Kiryandongo* [Settlement Profile]. UNHCR. https://data2.unhcr.org/en/documents/details/70207

United Nations High Commissioner for Refugees (UNHCR). (2021). *Global Trends: Forced Displacement in 2021*. https://www.unhcr.org/media/40152

Uppsala Conflict Data Program. (n.d.-a). CNDP. In *UCDP Conflict Encyclopedia*. Retrieved 5 February 2023, from https://ucdp.uu.se/actor/426

Uppsala Conflict Data Program. (n.d.-b). DR Congo (Zaire): Government. In *UCDP Conflict Encyclopedia*. Uppsala University. Retrieved 5 January 2023, from https://ucdp.uu.se/conflict/283

Uslaner, E. M. (2002). *The Moral Foundations of Trust*. Cambridge University Press.

Uslaner, E. M. (2008). Where You Stand Depends Upon Where Your Grandparents Sat: The Inheritability of Generalized Trust. *Public Opinion Quarterly*, *72*(4), 725–740. https://doi.org/10.1093/poq/nfn058

Verney, M.-H., & Clark, J. (2005, April 5). Tanzania camp to close as Burundian refugees return home. *UNHCR*. https://www.unhcr.org/my/news/stories/tanzania-camp-close-burundian-refugees-return-home

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P., & Girardin, L. (2015). Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family. *Journal of Conflict Resolution*, *59*(7), 1327–1342. https://doi.org/10.1177/0022002715591215

Walklate, S. (1998). Crime and Community: Fear or Trust? *The British Journal of Sociology*, *49*(4), 550. https://doi.org/10.2307/591288

Whiteley, P. F. (2000). Economic Growth and Social Capital. *Political Studies*, *48*(3), 443–466. https://doi.org/10.1111/1467-9248.00269

Wishman, M., & Butcher, C. (2022). Beyond ethnicity: Historical states and modern conflict. *European Journal of International Relations*, *28*(4), 777–807. https://doi.org/10.1177/13540661221106911

Wooldridge, J. M. (2007). *What's New in Econometrics? Lecture 10 Difference-in-Differences Estimation*. NBER Summer Institute. https://users.nber.org/~confer/2007/si2007/WNE/Slides7-31-07/slides_10_diffindiffs.pdf

Zak, P. J., & Knack, S. (2001). Trust and Growth. *The Economic Journal*, *111*(470), 295–321. https://doi.org/10.1111/1468-0297.00609

Zhou, Y.-Y., & Shaver, A. (2021). Reexamining the Effect of Refugees on Civil Conflict: A Global Subnational Analysis. *American Political Science Review*, *115*(4), 1175–1196. https://doi.org/10.1017/S0003055421000502

Zmerli, S., & Newton, K. (2008). Social Trust and Attitudes Toward Democracy. *Public Opinion Quarterly*, *72*(4), 706–724. https://doi.org/10.1093/poq/nfn054

Appendix

Appendix A. Supplementary material.

Afrobarometer sampling, geocoding, and GADM cross-referencing.

The geocoding by AidData follows a double-blind coding system to spatially identify the "Enumeration Areas" where respondents are interviewed. The Enumeration Area is a core feature of the Afrobarometer sampling strategy to achieve representativeness within countries. Sampling is done in five random sampling stages, where starting point, household, and respondent are drawn randomly (Afrobarometer, n.d.). However, for each Enumeration Area, only eight interviews are conducted to keep logistical fieldwork limited, i.e., the interviews are clustered. The methods section explains how the logistic analysis is performed with robust standard errors clustered on the Enumeration Area.

What AidData does post-survey is identify the coordinates for each enumeration area, where longitudinal and latitudinal coordinates for each respondent is provided. Two geocoders independently use a series of databases to establish the precise coordinates of an Enumeration Area based on Afrobarometer place names. The methodology is more accurately described in (BenYishay et al., 2017). Further quality assurances are conducted on random samples of the dataset where expert coders tested if it was possible to achieve higher granularity. For the whole of the Afrobarometer dataset, round 3 (5) 3467 (6649) unique Enumeration Areas were identified and given longitudinal and latitudinal coordinates. Location names and administrative units are also included in the PRIO Afrobarometer geocoded datasets, building on the geocoded AidData Afrobarometer datasets. The location names and administrative units are however not consistent across rounds. To overcome this, the Afrobarometer round 3 and 5 GADM coding in the Ethnic Stratification (ES) dataset, provided by Hodler et al., is utilised (2020).

To ensure consistency In the ES dataset, GADM codes have been coded, but not the location name, which is necessary for identifying where respondents live in conjecture with refugee settlements. To assure that both the GADM coding is correct, and ascertain the location name, I cross-reference the GADM codes and names in the GADM shapefile in QGIS with the ES GADM codes and the geocoded PRIO Afrobarometer dataset.

The regions coded in the PRIO Geocoded Afrobarometer dataset do not correspond 1:1 with the GADM 4.1 shapefile level 1 place names. Some respondents are ascribed a larger region. The greatest frequency discrepancies often correspond well to the GADM locations that are missing location names in the PRIO Afrobarometer set. To minimize the degree to which this is guesswork, I use the GADM 4.1 database and shapefile in QGIS to visually verify that the missing location is in fact an adjacent GADM unit to the location with a similar, but opposite, frequency discrepancy. In most of these instances, the discrepancy can be explained by a missing location 100%, but not always. The most noteworthy here is the Northern discrepancy in Zambia. In the ES dataset, ZMB.8_1 (which is Northern in the GADM shapefile) have 72 respondents fewer than "Northern" location name in PRIO Afrobarometer. "Muchinga" is not a region in the PRIO Afrobarometer dataset round 3, but a part of GADM4.1 and the ES dataset, with 88 respondents. Combined with the "Eastern" discrepancy and adjacent position of the three regions, I believe the discrepancy of respondents that registers as "Northern" in PRIO Afrobarometer to be part of "Muchinga" in the ES dataset. This is the most glaring example, yet it also describes my methodology for understanding, mapping, and coding discrepancies. Given the rather meticulous checking of the GADM coding, I trust that the other discrepancies of 8-16 respondents are due to similar cases. Although the ES dataset and GADM is consistent across rounds 3 and 5, the geocoded Afrobarometer placenames are not. E.g., in Afrobarometer round 5, "Muchinga" is included as a placename, but not in round 3. Thus, I do the GADM label correction process on split ES and AB samples, once for round 3 and once for round 5, before remerging the datasets and correctly label the GADM units.

Appendix B. Descriptive statistics for split country samples

Table B1. Descriptive Statistics Uganda (N = 4735)

	Freq.	Mean	SD	Min.	Max.	Median
Generalised social trust		.16	.37	.00	1.00	.00
Round (dummy)						
Round 3	2377	.50				
Round 5	2358	.49				
Admin1 refugee settlement any						
No settlement	3894	.82				
Any settlement	841	.17				
Admin1 refugee settlement categorical						
No settlement	3926	.82				
DRC settlement	311	.07				
Other settlement	498	.11				
Distance matrix other refugee settlements 2005/2010 (logged)		11.87	.68	8.39	12.75	12.10
Distance matrix DRC refugee settlements 2005/2010 (logged)		11.98	.72	8.58	13.03	12.12
Male (dummy)		.50	.50	.00	1.00	1.00
Age		33.88	12.58	18.00	88.00	30.00
Present living conditions		2.53	1.25	1.00	5.00	2.00
Enumeration Area: Health clinic (dummy)		.72	.45	.00	1.00	1.00
Enumeration Area: Electrical grid (dummy		.39	.49	.00	1.00	.00
Enumeration Area: Piped water system (dummy)		.30	.46	.00	1.00	.00
Education categorical						
Less than primary schooling	1713	.36				
Primary schooling	1931	.41				
Secondary schooling	567	.12				
Post secondary schooling	524	.11				

Table B2. Descriptive Statistics Tanzania (N = 3596)

	Freq.	Mean	SD	Min.	Max.	Median
Generalised social trust		.12	.33	.00	1.00	.00
Round (dummy)						
Round 3	1217	.33				
Round 5	2379	.66				
Admin1 refugee settlement any						
No settlement	3139	.87				
Any settlement	457	.12				
Admin1 refugee settlement categorical						
No settlement	3139	.87				
DRC settlement	134	.04				
Other settlement	323	.09				
Distance matrix other refugee settlements 2005/2010 (logged)		12.80	.94	5.82	13.94	12.92
Distance matrix DRC refugee settlements 2005/2010 (logged)		13.17	.76	9.56	14.05	13.38
Male (dummy)		.50	.50	.00	1.00	1.00
Age		37.96	13.86	18.00	99.00	35.00
Present living conditions		2.32	.98	1.00	5.00	2.00
Enumeration Area: Health clinic (dummy)		.39	.49	.00	1.00	.00
Enumeration Area: Electrical grid (dummy		.29	.46	.00	1.00	.00
Enumeration Area: Piped water system (dummy)		.44	.50	.00	1.00	.00
Education categorical						
Less than primary schooling	855	.24				
Primary schooling	2369	.66				
Secondary schooling	276	.08				
Post secondary schooling	96	.03				

Table B4. Descriptive Statistics Zambia (N = 2187)

	Freq.	Mean	SD	Min.	Max.	Median
Generalised social trust		.11	.31	.00	1.00	.00
Round (dummy)						
Round 3	1101	.51				
Round 5	1086	.49				
Admin1 refugee settlement any						
No settlement	1529	.69				
Any settlement	658	.30				
Admin1 refugee settlement categorical						
No settlement	1529	.70				
DRC settlement	308	.14				
Other settlement	350	.16				
Distance matrix other refugee settlements 2005/2010 (logged)		12.79	.75	6.50	13.77	12.98
Distance matrix DRC refugee settlements 2005/2010 (logged)		12.75	.69	9.89	13.59	12.92
Male (dummy)		.51	.50	.00	1.00	1.00
Age		34.17	13.06	18.00	90.00	31.00
Present living conditions		2.85	1.22	1.00	5.00	3.00
Enumeration Area: Health clinic (dummy)		.58	.49	.00	1.00	1.00
Enumeration Area: Electrical grid (dummy		.53	.50	.00	1.00	1.00
Enumeration Area: Piped water system (dummy)		.36	.48	.00	1.00	.00
Education categorical						
Less than primary schooling	487	.22				
Primary schooling	916	.42				
Secondary schooling	458	.21				
Post secondary schooling	326	.15				

Appendix C. Pairwise correlation matrix for full sample

Table C1. Pairwise correlations matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) Generalised social trust	1.000												
(2) Round (dummy)	-0.009	1.000											
(3) Admin1 refugee settlement dummy	0.038	-0.042	1.000										
(4) Admin1 refugee settlement categorical	0.033	-0.035	0.938	1.000									
(5) Distance matrix other refugee settlements 2005/2010 (logged)	-0.057	0.282	-0.409	-0.442	1.000								
(6) Distance matrix DRC refugee settlements 2005/2010 (logged)	-0.034	0.095	-0.356	-0.236	0.465	1.000							
(7) Male (dummy)	0.006	-0.003	0.003	0.003	-0.003	0.002	1.000						
(8) Age	0.026	0.101	0.011	0.016	0.058	0.090	0.146	1.000					
(9) Present living conditions	0.003	-0.108	0.090	0.098	-0.022	-0.058	-0.019	-0.132	1.000				
(10) Enumeration Area: Health clinic	0.006	0.031	0.043	0.028	-0.157	-0.149	-0.004	-0.056	0.026	1.000			
(11) Enumeration Area: Electrical grid	-0.076	-0.026	-0.136	-0.122	0.067	0.022	-0.007	-0.088	0.080	0.258	1.000		
(12) Enumeration Area: Piped water system	-0.063	0.066	-0.098	-0.095	0.110	0.109	-0.004	-0.048	0.085	0.132	0.517	1.000	
(13) Education categorical	-0.076	-0.041	-0.008	-0.001	0.026	0.023	0.115	-0.174	0.148	0.109	0.261	0.234	1.000

Appendix D. Logistic models for split country samples

Table D1. Part II logistic models. Uganda

VARIABLES	(1) Trust	(2) Trust	(3) Trust	(4) Trust	(5) Trust	(6) Trust
Round 5 (dummy)	0.001	7.436***	-11.666***	-0.036	8.933***	-10.947***
	(0.107)	(1.519)	(1.614)	(0.103)	(1.460)	(1.609)
Admin1 refugee settlement						
DRC settlement	0.874***			0.785***		
	(0.215)			(0.223)		
Other settlement	0.087			-0.043		
	(0.203)			(0.196)		
Round 5 × Admin1 DRC settlement	-1.007***			-1.034***		
	(0.334)			(0.355)		
Round 5 × Admin1 other settlement	0.328			0.469		
	(0.307)			(0.303)		
Distance matrix other settlement (logged)		0.132*			0.236***	
		(0.079)			(0.075)	
Round 5 × Distance matrix other settlement		-0.629***			-0.758***	
		(0.128)			(0.123)	
Distance matrix DRC settlement (logged)			-0.478***			-0.430***
			(0.078)			(0.074)
Round 5 × Distance matrix DRC settlement			0.970***			0.909***
			(0.134)			(0.134)
Male (dummy)				0.060	0.059	0.061
				(0.073)	(0.073)	(0.073)
Age				0.005*	0.006*	0.006*
				(0.003)	(0.003)	(0.003)
Present living conditions				0.004	-0.005	0.021
				(0.033)	(0.034)	(0.033)
Enumeration Area: Health clinic				0.177	0.249**	0.055
				(0.125)	(0.119)	(0.120)
Enumeration Area: Electrical grid				-0.284**	-0.341**	-0.282**
				(0.144)	(0.147)	(0.142)
Enumeration Area: Piped water				-0.188	-0.179	-0.134
				(0.161)	(0.165)	(0.163)
Education.						
Primary schooling				-0.160*	-0.160	-0.134
				(0.097)	(0.097)	(0.097)
Secondary schooling				-0.067	-0.043	-0.088
				(0.142)	(0.142)	(0.143)
Post secondary schooling				-0.742***	-0.762***	-0.735***
				(0.185)	(0.186)	(0.184)
Constant	-1.692***	-3.155***	4.066***	-1.723***	-4.456***	3.454***
	(0.077)	(0.921)	(0.924)	(0.190)	(0.896)	(0.879)
Observations	4,735	4,735	4,735	4,735	4,735	4,735

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.10

Table D2. Part II logistic models. Tanzania

VARIABLES	(1) Trust	(2) Trust	(3) Trust	(4) Trust	(5) Trust	(6) Trust
Round 5 (dummy)	-0.080	0.234	-0.624	-0.026	0.107	-0.158
	(0.132)	(1.707)	(1.887)	(0.139)	(1.629)	(1.908)
Admin1 refugee settlement						
DRC settlement	-0.677*			-0.708**		
	(0.376)			(0.350)		
Other settlement	-0.281			-0.320		
	(0.314)			(0.292)		
Round 5 × Admin1 DRC settlement	-0.607			-0.689		
	(0.799)			(0.792)		
Round 5 × Admin1 other settlement	-0.127			-0.146		
	(0.402)			(0.389)		
Distance matrix other settlement (logged)		0.172*			0.177*	
		(0.098)			(0.091)	
Round 5 × Distance matrix other settlement		-0.035			-0.022	
		(0.134)			(0.128)	
Distance matrix DRC settlement (logged)			0.096			0.147
			(0.111)			(0.112)
Round 5 × Distance matrix DRC settlement			0.040			0.009
			(0.144)			(0.146)
Male (dummy)				0.003	0.005	0.004
				(0.084)	(0.084)	(0.084)
Age				0.001	0.001	0.001
				(0.004)	(0.004)	(0.004)
Present living conditions				0.099	0.093	0.092
				(0.065)	(0.064)	(0.064)
Enumeration Area: Health clinic				-0.026	-0.011	-0.028
				(0.122)	(0.123)	(0.122)
Enumeration Area: Electrical grid				-0.287*	-0.267*	-0.276*
				(0.149)	(0.148)	(0.150)
Enumeration Area: Piped water				0.054	0.053	0.018
				(0.125)	(0.124)	(0.122)
Education.						
Primary schooling				-0.129	-0.143	-0.135
				(0.122)	(0.121)	(0.122)
Secondary schooling				-0.662**	-0.658**	-0.650**
				(0.263)	(0.264)	(0.263)
Post secondary schooling				-0.525	-0.546	-0.527
				(0.360)	(0.361)	(0.360)
Constant	-1.869***	-4.039***	-3.188**	-1.980***	-4.198***	-3.929***
	(0.107)	(1.205)	(1.448)	(0.290)	(1.163)	(1.451)
Observations	3,596	3,596	3,596	3,596	3,596	3,596

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.10

Table D3. Part II logistic models. Zambia

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Trust	Trust	Trust	Trust	Trust	Trust
Round 5 (dummy)	0.550**	1.484	-20.691***	0.453*	2.066	-17.890***
	(0.238)	(2.872)	(3.979)	(0.251)	(2.449)	(3.588)
Admin1 refugee settlement						
DRC settlement	1.499***			1.377***		
	(0.302)			(0.314)		
Other settlement	1.017***			0.955***		
	(0.295)			(0.349)		
Round 5 × Admin1 DRC settlement	-2.230***			-2.288***		
	(0.458)			(0.473)		
Round 5 × Admin1 other settlement	0.012			-0.089		
	(0.376)			(0.416)		
Distance matrix other settlement (logged)		-0.289**			-0.268*	
		(0.125)			(0.144)	
Round 5 × Distance matrix other settlement		-0.096			-0.151	
		(0.226)			(0.194)	
Distance matrix DRC settlement (logged)			-0.580***			-0.546***
			(0.133)			(0.152)
Round 5 × Distance matrix DRC settlement			1.631***			1.402***
			(0.310)			(0.280)
Male (dummy)				0.140	0.146	0.133
				(0.125)	(0.122)	(0.126)
Age				0.006	0.006	0.008
				(0.005)	(0.005)	(0.005)
Present living conditions				0.079	0.077	0.087
				(0.061)	(0.060)	(0.061)
Enumeration Area: Health clinic				-0.114	-0.035	0.162
				(0.189)	(0.181)	(0.182)
Enumeration Area: Electrical grid				-0.406*	-0.435**	-0.519**
				(0.225)	(0.219)	(0.233)
Enumeration Area: Piped water				-0.185	-0.363	-0.196
				(0.233)	(0.236)	(0.246)
Education.						
Primary schooling				-0.194	-0.209	-0.166
				(0.170)	(0.177)	(0.177)
Secondary schooling				-0.593**	-0.613***	-0.503**
				(0.233)	(0.232)	(0.228)
Post secondary schooling				-0.460*	-0.432	-0.472*
				(0.267)	(0.266)	(0.265)
Constant	-2.763***	1.415	5.081***	-2.618***	1.281	4.545**
	(0.150)	(1.542)	(1.676)	(0.305)	(1.838)	(1.944)
Observations	2,187	2,187	2,187	2,187	2,187	2,187

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.10

Appendix E. DDD estimate for split country samples

Table E1: DDD predicted probabilities Uganda

	Round 3	Round 5	DD(D) ATE
No settlement group	.1599	.1552	
Other settlement group	.1543	.2185	.0689
DRC settlement group	.2919	.1255	-.2353

Table E2: DDD predicted probabilities Tanzania

	Round 3	Round 5	DD(D) ATE
No settlement group	.1305	.1276	
Other settlement group	.0984	.0842	-.0113
DRC settlement group	.0690	.0351	-.0227

Table E3: DDD predicted probabilities Zambia

	Round 3	Round 5	DD(D) ATE
No settlement group	.0650	.0981	
Other settlement group	.1510	.2023	.0183
DRC settlement group	.2116	.0423	-.1876

NTNU

Kunnskap for en bedre verden