THE DARFUR GENOCIDE:
Comparing Alternative Explanations with Geographic Information Systems

Coen Ottevanger – s1085409

7 July 2017

Universiteit Leiden

with supervision of :
Dr. Eelco van der Maat

Word count :
15.291

This thesis is submitted to the Faculty of Humanities, Leiden University, in partial fulfilment of the requirement for the degree of Master of Arts International Relations. Please use a digital version of this thesis for the depiction of colours.

LISTS OF VARIABLES, SUBSCRIPTS AND ABBREVIATIONS

Variables	Description	Units
A	Surface area	m^2
BWS	Baseline Water Stress	-
D_{FMZ}	Inverted distance from centroids of counties to centroids of ethnic territories	m^{-1}
DSI	Drought Severity Index	-
E_{FMZ}	Population density of Fur, Massaleit and Zaghawa tribes	m^{-2}
M	Original DSI in raw data	-
N	Sample size	-
N_{oil}	Number of oil production sites	-
P_{oil}	Percentage of the county overlapping with oil concession areas	-
Pop	Population	-
r	Pearson correlation coefficient	-
R^2	Coefficient of determination	-
SCI	Soil Competition Index	m^{-2}
y	Dependent variable (measure for the severity of violence)	-
$Z_{comparison}$	Z-Scores for comparing Pearson correlation coefficients	-
β	Regression coefficient	-
φ	Percentage of destroyed structures per village	-

Subscripts	Description
FMZ	of Fur, Massaleit and Zaghawa origin
j	Village
m	Collection of villages in county i
i	County

Abbreviation	Description
AMIS	African Union Mission in Sudan
DPA	Darfur Peace Agreement
FMZ	Fur, Massaleit and Zaghawa
GIS	Geographic information system
ICC	International Criminal Court
ICID	International Committee of Inquiry into Darfur
JEM	Justice and Equality Movement
SLA	Sudan Liberation Army
UNSC	United Nations Security Council
VIF	Variance Inflation Factor

CONTENTS

1. INTRODUCTION

Far from Sudan's heartland on the Nile and only linked to the country's capital Khartoum with an endless dirt road lies the region Darfur, which literally translates as 'home of the Fur.' The deserts of North Darfur, the hills of East Darfur, the *wadis* of West Darfur and the mountain massif of Central Darfur conjointly approximate the size of Spain, but only inhabited around a tenth of Spain's population in 2003; four to six million people. The vast majority of them are Muslim, but the province is a mosaic of Arab and non-Arab ethnic groups.[1] The Fur are, together with the Massaleit and Zaghawa, the largest non-Arab groups. Besides these three groups, there are relatively newer non-Arab tribes, who have been neutral in the recent conflict.[2] Some lines between the Arab and non-Arab tribes blurred. Although many of the non-Arab groups have their own indigenous language, many speak Arabic too.[3] While their main occupation is sedentary agriculture and the majority of Arab groups herd cattle and are nomadic, over the years many Fur have also acquired cattle. In addition, the skin tone of the non-Arab and Arab tribes are indistinguishable and individuals can often switch tribes via marriage in just a single generation.

Background

This ethnic mix of people has faced many political problems since its independence from the United Kingdom and Egypt in 1956. Two civil wars between the central government in Khartoum and separatists in the south of Sudan devastated the country between 1955 and 1972 *and* between 1983 and 2005, eventually leading to the secession of South Sudan from Sudan in 2011. In the midst of this violence, Omar al-Bashir rose to power through a coup in 1989 and remained in power ever since, ruling autocratically with a strong underpinning of Islamism and Arab supremacy, which has brutally excluded non-Arab Darfurians from roles in the regime.

Darfur always had a troubled relationship with Sudan and its central government in Khartoum. After an autonomous period as a sultanate from 1600, Darfur was incorporated into Sudan in 1916. Soon afterwards, Darfur was increasingly marginalised and exploited.[4] This marginalisation created tensions between ethnic groups in Darfur. In the 1980s already, droughts and the resulting famine had fuelled violence between Fur and Arab groups in Darfur

[1] A term coined by: Marco Boggero, "Darfur and Chad: A Fragmented Ethnic Mosaic," *Journal of Contemporary African Studies* 27, no. 1 (2009).
[2] Ola Olsson and Michele Valsecchi, "Quantifying Ethnic Cleansing: An Application to Darfur," (2010): 11.
[3] See: Heather J Sharkey, "Arab Identity and Ideology in Sudan: The Politics of Language, Ethnicity, and Race," *African Affairs* 107, no. 426 (2008).
[4] John Hagan and Joshua Kaiser, "The Displaced and Dispossessed of Darfur: Explaining the Sources of a Continuing State- Led Genocide," *The British Journal of Sociology* 62, no. 1 (2011).

as migration fluxes into fertile areas escalated resource scarcity.[5] The violence intensified further as Gaddafi's Libya supplied arms to Arab herders in Darfur to advance pan-Arabism and Islamism into Sudan.[6] With this conflict and with later marginalisation, the ethnic mosaic increasingly turned into an Arab versus non-Arab dichotomy.

The conflicts between Fur and Arab tribes never fully resolved. Low-level warfare remained endemic to the region until it escalated into another tribal war between Massaleit and Arab tribes from 1995 to 1999. Meanwhile, anxiety amongst the non-Arabs grew further as Arab militias gained political influence and impunity.[7] In 2003, this low-level warfare intensified again towards a full-scale war and genocide in Darfur.

Most narratives of the war in Darfur start with the attack by the Sudan Liberation Army (SLA) and the Justice and Equality Movement (JEM) on an airport at al Fashar in April 2003, where the SLA and JEM had secured weapons and vehicles for the violence to come.[8] A few months after this attack, the violence escalated further as Omar al-Bashir ordered attacks on the region as a response to the rebel's attack. Government officials urged the Janjaweed, a militia that mainly recruits from Arab tribes and that had received al-Bashir's support in the forms of arms and money before,[9] to increase the recruitment of militiamen from June 2003.[10] After this public endorsement, Musa Hilal, the leader of the Janjaweed, repeatedly expressed his intent to eliminate all non-Arabs from Darfur.[11] The tactics were clear, Sudanese aircraft bombings preceded the Janjaweed's rape, loot, killings and destructions. Between February 2003 and April 2006 alone, the violence killed around 235.000 people. Indirect damage, through starvation and diarrhoea, killed another 200.000 people at the same time.[12] Additionally, over six million people fled their homes. The majority of refugees has no prospect of returning to their villages and their refugee camps have slowly transformed into permanent cities as the war continues.[13]

The international community's response was inadequate. The African Union Mission in Sudan (AMIS) in April 2005 and the Darfur Peace Agreement (DPA) of 2006 both failed to de-escalate the conflict. AMIS could not offer Darfurians protection. The Sudanese bombing of two rebel villages, a few days after the head of the African Union's peacekeeping force had visited,

---

[5] See: Alex De Waal, *Famine That Kills: Darfur, Sudan* (New York: Oxford University Press 1989).
[6] Hagan and Kaiser, 5.
[7] John Ryle, "Disaster in Darfur," *New York Review of Books* 51, no. 13 (2004).
[8] Jess Gifkins, "Darfur," in *The Oxford Handbook of the Responsibility to Protect*, ed. Alex Bellamy and Tim Dunne (Oxford: Oxford University Press, 2016), 718.
[9] The evidence for ties between Janjaweed and the Sudanese government are irrefutable and summarised in several reports of the Human Rights Watch, for the best report, see: "Entrenching Impunity: Government Responsibility for International Crimes in Darfur," (Human Rights Watch, 2005).
[10] Hagan and Kaiser, 6.
[11] John Hagan and Wenona Rymond-Richmond, *Darfur and the Crime of Genocide* (New York: Cambridge University Press, 2008).
[12] Olivier Degomme and Debarati Guha-Sapir, "Patterns of Mortality Rates in Darfur Conflict," *The Lancet* 375 (2010): 295.
[13] Paul Jeffrey and Annie Bungeroth, "In Pictures: Darfur Refugees Then and Now," BBC, http://www.bbc.com/news/world-africa-28739000.

exemplifies this inability and, additionally, aggravated the distrust between Darfurians and foreign military forces.[14]

The underlying reason for the late response of the international community was that the United Nations Security Council (UNSC) was held in a stalemate as China only voted for resolutions that required Sudanese consent. In August 2006, the UNSC adopted Resolution 1706. The resolution would enforce the DPA via deployment of 22.500 peace-support personnel to Darfur. However, eight months later, only 200 technical personnel were deployed. The UNSC Resolution 1706 basically gave Khartoum's regime a veto power over the deployment of a force charged with halting the genocide executed by that same regime.[15] China's U-turn only occurred as the 2008 Summer Olympic Games in Beijing approached.[16] International advocacy persuaded China to vote in favour of UNSC Resolution 1769, which fully authorised the deployed of 23.000 personnel to Darfur from August 2007 onwards. Although violence was impeded to some extent and an important peace agreement was signed in May 2011, violence often reawakens and continues to destroy lives until this day.

Outline

In the light of the atrocities in Darfur, scholars have suggested several causes of the violence. Scholars have focused on four theories to explain the violence in Darfur: ethnicity, resource scarcity, counter-insurgency and intra-elite rivalry. But until now, it remains unclear which explanation has more explanatory power. This thesis aims to shed light on the relative importance of two of these alternative explanations, via a statistical analysis of geo-coded data on Darfur's counties using a geographic information system (GIS) and a unique dataset compiled out of a large variety of publicly available data. On the basis of a thorough scan of available data, this thesis focuses solely on ethnic tension and resource scarcity as explanations for the violence in Darfur. Simply put, the main research question is: which of two explanations, ethnicity or resource scarcity, contributed more to the root causes of the violence in Darfur?

[14] See: Willa Thayer, "Sudan Bombs Rebel Areas in Darfur after Au Peacekeeper Visit," the Independent, http://www.independent.co.uk/news/world/africa/sudan-bombs-rebel-areas-in-darfur-after-au-peacekeeper-visit-430459.html.
[15] Eric Reeves, "Getting Darfur Wrong," *Dissent* 56, no. 4 (2009).
[16] Ibid.

To answer this question, two linear regression models are designed. The performed linear regressions statistically correlate quantifiable and observable implications of the two alternative explanations to the dependent variable: the spatial variation of the severity of violence within Darfur. The statistical units of this data analysis are the sixty-three counties of Darfur. The accuracy and predictive strength of each model quantifies the explanatory power of the two theories, thereby answering the main question.

Concerning the two explanations, the reading of literature puts forward four hypotheses: the severity of violence increases with (i) the population density of Fur, Massaleit and Zaghawa tribes in each county, (ii) with the proximity of counties to the centre of ethnic territories, (iii) with the resource scarcity in each county and (iv) with the presence of oil in each county. The first two hypotheses (i and ii) fall under the umbrella of the explanation regarding ethnic tensions while the last two (iii and iv) relate to the view that the Darfur genocide was primarily a struggle over resources.

The wide variety of used data and the chosen geographical approach makes this research a fresh addition to existing literature. No other empirical research has attempted to analyse the entire area of Darfur while simultaneously shedding light on different explanations.[17] In addition, no other study uses micro-level data to describe the region. Finally, in literature, qualitative methods are primarily compared rival explanations while statistical methods primarily verified the validity of single explanations. Therefore, the application of statistical methods to compare alternative explanations offers a unique opportunity. This wide scope might elucidate previously unobservable patterns.

[17] Ola Olsson tested ethnic tensions and resource scarcity with a more limited set of data that also only described in around five percent of Darfur: Ola Olsson and Eyerusalem Siba, "Ethnic Cleansing or Resource Struggle in Darfur? An Empirical Analysis," *Journal of Development Economics* 103 (2013).

Many have seen the relevance of studying genocides before this research; the war in Darfur is the subject of hundreds of books, speeches, reviews and news articles. A better understanding of the causes and dynamics of genocide is relevant in several ways. Firstly, it can help prevent future atrocities by identifying risk factors and indicators early on. Thereby the case of Darfur can be an important lesson for the international community on how (not) to act. Additionally, during the international response to war, understanding the process and causes of genocide can help humanitarian organisations and peace-keeping missions to target their aid effectively and safely. Furthermore, this research produces a visually appealing way to explain causes of the violence in Darfur, which could be used as a tool for educational purposes. Finally, clearly identified patterns of genocide can serve as paramount evidence for a more widely accepted declaration of genocide and prosecution of its culprits.

The remainder of this paper first elaborates on all the alternative explanations in academics. A preliminary examination of available data then shows that only the influence of ethnic tensions and resource scarcity on violence is suitable for testing. The chapter on research design then elaborates further on how these explanations are tested and how these tests answer the research question. The research design also identifies the observable implications of each theory and how these implications quantified into analysable variables. It also addresses the scope of the research, the statistical strategy and all its limitations. The final chapters analyse the results of the designed linear regressions and draw conclusion from the observations.

2. LITERATURE REVIEW

The Darfur genocide has the perverse honour of being the most fully chronicled genocide since the Holocaust. The hundreds of human rights reports, newspaper articles, governmental investigations, accounts of non-governmental agencies, reports of aid workers, research papers and policy recommendations show a relative conformity around key themes.

Although it has taken years to form this consensus and there still are many dissenting voices,[18] there is now a broad academic agreement about the role of the Sudanese government in the deaths of hundreds of thousands and the displacement of six million Darfurians, [19] overwhelmingly from the non-Arab tribal populations of the region. This consensus further strengthened in July 2010 through the International Criminal Court's (ICC) arrest warrant for Omar al-Bashir on the count of genocide. However, the ICC's warrant failed to solve Darfur's problem as it demonised the regime so that agreements between the regime and rebels became virtually impossible.[20] Despite dissenting voices, this thesis refrains from joining the legal, and often semantic, debate on the use of term 'genocide' and defines the atrocities in Darfur as genocide.[21]

Apart from this debate on Darfur's genocide, the causes of this genocide are heavily debated too. Four explanatory factors dominate the academic debate: ethnicity, resource scarcity, counter-insurgency and intra-elite rivalry. As this reading of the literature argues, all the explanations roughly fit into a single narrative. The explanations that fit into this narrative do not inherently contradict their alternatives; they are simply true on different levels. For example, ethnicity can co-exist with, or even have a correlation to the rebellion. The Darfur genocide is, therefore, a typical case of overdetermination. This overdetermination also creates the confusion and the difficulties in determining the relative importance of alternative explanations. The following literature review aims to untangle explanations into clear categories.

[18] The U.N. and the African Union for instance still refrain from using the term 'genocide' as prove for 'genocidal intent' was limited in their eyes.
[19] Reeves.
[20] Martin Shaw, "Darfur: Counter- Insurgency, Forced Displacement and Genocide," *The British Journal of Sociology* 62, no. 1 (2011).
[21] The genocide debate focuses on if genocide must originate from a genocidal intent, or if genocide can also occur as a side effect of a counter-insurgency. The debate has been further characterised by idea that genocide *per se* entails mass murder and not forced displacement, see: Alex De Waal, "Counter-Insurgency on the Cheap," *Review of African Political Economy* 31, no. 102 (2004); Hagan and Kaiser. For some more reflections on the genocide debate, see: SMH Jennifer Leaning MD, "Diagnosing Genocide-the Case of Darfur," *The New England journal of medicine* 351, no. 8 (2004); David Luban, "Calling Genocide by Its Rightful Name: Lemkin's Word, Darfur, and the Un Report," *Chicago Journal of International Law* 7, no. 1 (2006); Scott Straus, "Darfur and the Genocide Debate," *Foreign Affairs* 84, no. 1 (2005); Mika Vehnämäki, "Darfur Scorched: Looming Genocide in Western Sudan " *Journal of Genocide Research* 8, no. 1 (2006).

Ethnicity

Firstly, ideology would explain why the never-ending marginalisation of Darfurians, as much of Darfur's population differs ethnically from the Arab regime in Khartoum. Because Sudanese and Darfurians are all Muslim, it is not a religion but the ethnicity based on Arabism that explains the centuries-old tensions and the current conflict. Ryle argued that the internal history of conflict "between ethnic groups in Darfur is one of Balkan complexity."[22] Scholars subscribe to this conclusion, referring to historic ethnicity in Darfur as a mosaic rather than a binary polarised society between Arabs and Africans.[23]

The explanation of ethnicity works on different levels as Arabism has affected both the beliefs of the national leader, Omar al-Bashir, and the interaction between tribes within Darfur. Additionally, ethnicity plays a role in the root causes of the conflict and in the rollout of the actual violence.

Ethnic divisions ignited centuries ago, as *Ta'rib* or Arabisation gradually spread through Sudan, slowly propagating the Arab identity and language into remote regions as Darfur. This Arab supremacist ideology spread further during Gaddafi's campaigns to unify Libya, Chad and Sudan in the 1980s.[24] In close collaboration with the Arab Um Julal tribe, this Libyan campaign created the Arab Gathering, an "Islamist militantly racist and pan-Arabist organisation that later supplied the founders of the Janjaweed."[25] The Um Julal again are a part of the Abbala (camel-herding) Rizeigat tribe, whose quarrels with the Fur date back to the Fur Sultanate of the seventeenth century when.[26] The old Sultan divided Darfur's land. Known as *hawakir*, these territorial grants are the basis of today's ethnic land divisions. The Abbala Rizeigat, the ethnic base of Janjaweed, still ascribes their role in the conflict to them not receiving territory over three hundred years ago.[27] This shows the relevance of Darfur's ethnic history in the recent genocide.

Ethnicity and Arabism also plays a role on a larger scale. Sudan's government has consequently marginalised the population of Darfur along ethnic lines. Since his coup in 1989, Omar al-Bashir has mostly appointed Arab officials. The government often altered borders to make the

---

[22] Ryle.
[23] Alex De Waal, "Deep Down in Darfur: Nothing Is as We Are Told in Sudan's Killing Fields," review of Darfur: the Ambigious Genocide by Gerard Prunier, *Review of African Political Economy* 32, no. 106 (2005): 658.
[24] Julie Flint and Alex De Waal, *Darfur: A New History of a Long War*, ed. Julie Flint and Alex De Waal, African Arguments (London: Zed Books, 2008), 47 - 50.
[25] Gerard Prunier, *Darfur: The Ambiguous Genocide* (Ithaca: Cornell University Press, 2005), 45.
[26] De Waal, "Deep Down in Darfur: Nothing Is as We Are Told in Sudan's Killing Fields," 654.
[27] Ibid.

Fur the minority in each Darfurian states,[28] which gave the ancient struggle for land and power an ethnic element.[29] Later, Khartoum even did away with hereditary rights to land and disarmed non-Arab villages, further unbalancing Darfur's Arab and non-Arab populations.[30] Medani specifically argued that these kinds of "Islamist-backed regime's policies in the 1990s ultimately led to the conflict in Darfur."[31] One manuscript that significantly contributed to the rebellion's appeal, *the Black Book*, clearly stipulates the effects that these policies had. *The Black Book* showed that Darfur scores far worse on representation, funding, health care, education and infant mortality than other regions do.[32] Apart from within the region, ethnicity and Arab supremacy thus also heavily influenced the behaviour of the regime in Khartoum towards non-Arab groups in Darfur.

The ethnicity thus played a strong role in the run-up to the war in Darfur, but it echoes through in the targets of the actual violence as well. Olsson and Valsecchi studied the violence in a small part of Darfur, accounting for 12 percent of Darfur's total population.[33] They found that 57 percent of Fur, Massaleit and Zaghawa were cleansed from their villages, indicating a rather 'successful' cleansing. In addition, 85 percent of the cleansed population also had their village destroyed. This violence caused the relative population of Fur, Massaleit and Zaghawa in that part of Darfur to drop from 70 percent to 47 percent in five years. Several other quantitative studies confirm these findings.[34] In addition, scholars have found that the Janjaweed consistently used racial epithets during the raids, often shouting, "We will kill any slaves we find and cut off their heads!" or "This is the last day for blacks!" and "You blacks are not human. We can do anything we want to you. You cannot live here!"[35] The statistics and the racial epithets show that that racism and ideology were not only at the root of the conflict but form an important motivation for individual practitioners of violence.

Overall, these regional and national dynamics also show how ideology intrinsically links to power-maximisation. In the regime's goal of annihilating the rebellion and further suppressing the Darfurians, the Janjaweed found a vicious opening to combine their own political and ethnic agenda with the agenda of the central government. Furthermore, the regime adheres to Arab supremacy in name, but Ryle pointed out it had already "purged itself of some of the hard-core

[28] Khalid Mustafa Medani, review of Darfur's Sorrow: A History of Destruction and Genocide (New York: Cambridge University Press, 2007), *International Journal of Middle East Studies* 42 no. 1 (2010): 175.
[29] Martin W Daly, *Darfur's Sorrow: A History of Destruction and Genocide* (New York: Cambridge University Press, 2007), 244.
[30] Reeves.
[31] Medani.
[32] Alex Cobham, "Causes of Conflict in Sudan: Testing the Black Book," *The European Journal of Development Research* 17, no. 3 (2005).
[33] Olsson and Valsecchi.
[34] John Hagan, Wemona Rymond-Richmond, and Patricia Parker, "The Criminology of Genocide: The Death and Rape of Darfur," *Criminology* 43, no. 3 (2005); Michael Vanrooyen et al., "Employment of a Livelihoods Analysis to Define Genocide in the Darfur Region of Sudan," *Journal of Genocide Research* 10, no. 3 (2008).
[35] John Hagan and Wenona Rymond-Richmond, "The Collective Dynamics of Racial Dehumanization and Genocidal Victimization in Darfur," *American Sociological Review* 73, no. 6 (2008): 882.

ideologues that had guided it until the late 1990s."[36] As Shaw concludes, apart from ethnic goals, the regional militia and the national leaders seemed "rational actors pursuing (inhuman) political goals."[37] Although there are many ideological and ethnic features to the conflict, it remains hard to distinguish an ideological agenda from a political one.

Resource scarcity

The second explanation, which received significant attention after a report of the United Nations Environment Programme and an essay by Jeffrey Sachs, proposes that climate change, land degradation and the resulting competition over natural resources as water and food is the root cause of the conflict.[38] Interestingly, the government of Sudan has consistently promoted this narrative as well.

The impact of changes in the water-food-climate nexus on the war in Darfur is twofold. Firstly, resource scarcity has started exacerbating ethnic tensions within Darfur decades before the conflict. Scholars argue that Arab nomads and non-Arab farmers had co-existed peacefully for years as land was abundant and local authorities adequately resolved conflicts.[39] Tensions grew after the Unregistered Land Act of 1970, which did away with the traditional land rule and put ownership of all unregistered land into the hands of the government. This meant that new tribes could settle anywhere while previously they were merely welcomed as guests of the hosting tribe.[40] This Unregistered Land Act thus especially benefited Arab nomads at the expense of sedentary, non-Arab landowners. The act also diminished the legitimacy of traditional courts, increasing the appeal of violence to resolve conflicts.

The pressure on land only grew further in the following years. In the 1980s, droughts, desertification and rapid population growth caused a famine that resulted in mass migration fluxes towards more fertile grounds. After the destructive famine, Arab tribesmen and semi-sedentary Fur communities violently clashed in an even more destructive war over resources.[41] The central government also used the resource scarcity to further marginalise Darfurians, as the destructive damages of droughts and famines could have been diminished with only minimal governmental competence in Khartoum.[42] According to Prunier, these were "the first signs of enmities between nomadic and semi-sedentary tribes in Darfur that set the stage for the violent

[36] Ryle, 57.
[37] Shaw, 60.
[38] Jeffrey D Sachs, "Ecology and Political Upheaval," *Scientific American* 295, no. 1 (2006).; "Sudan Post-Conflict Environmental Assessment ", (Nairobi: United Nations Environment Programme, 2007).
[39] Musa Adam Abdul-Jalil, "The Dynamics of Customary Land Tenure and Natural Resource Management in Darfur," *Land Reform FAO*, no. 2 (2006).
[40] Ibid., 18.
[41] Medani, 175.
[42] Jeffrey Kaplan, "A Review Of: Gerard Prunier. Darfur: The Ambigious Genocide"," *Terrorism and Political Violence* 20, no. 146 - 148 (2007).

conflict in Darfur."[43] Prunier added that the schism between nomadic and sedentary tribes was later superimposed on an Arab versus African dichotomy.[44] This shows how ethnic tensions and violence originate from resource scarcity.

The central government has also heavily invested in securing oil resources in the region. Although there is hardly any oil in production in Darfur, rumours about the widespread presence of oil have exacerbated tensions in the region.[45]

Another side of the explanation focuses on how capturing land and access to water played a role in the Janjaweed's rollout of the violence. Hagan and Kaiser argued that the chance to "take land from Black African farming groups played an explicit part in the mobilisation and training of the militias."[46] Vanrooyen added that after the Janjaweed had raided a village, "nothing remains on the surface to mark the boundaries of fields and generations of cultivation that defined the farm holdings,"[47] giving militiamen the opportunity to reclaim the land. He also showed that theft of livestock was a primary aim of the Janjaweed.[48] However, if scarcity contributed to the violence in Darfur and capturing resources was an aim of the Janjaweed, it is hard to explain why the Janjaweed deliberately poisoned wells, destroyed food stores and uprooted trees.[49]

Although many scholars have qualitatively tied resource scarcity to violence,[50] the statistical evidence for a link between resource scarcity and violence is inconclusive. One study found no evidence for a correlation between violence and changes in vegetation,[51] while another study did proof that same correlation.[52] Yet another study showed that trends in rainfall only weakly correlated with the outbreak of violence.[53] However, all these scholars fail to take into the full scope of elements of environmental degradation. Population growth (human and animal), changes in agriculture and cattle breeding, use of water for irrigation and ground water salinisation are all important, and neglected, aspects of environmental degradation. Collectively, these components might expose the correlation between environment and violence.

---

[43] Prunier.
[44] Ibid., 162.
[45] Andrea Behrends, "Fighting for Oil When There Is No Oil Yet: The Darfur-Chad Border," *European Journal of Anthropology* 2008, no. 52 (2008).
[46] Hagan and Kaiser, 7.
[47] Vanrooyen et al., 352.
[48] Ibid., 351.
[49] Hagan and Kaiser, 10.
[50] See for instance: Flint and De Waal.
[51] Ian Brown, "Assessing Eco-Scarcity as a Cause of the Outbreak of Conflict in Darfur: A Remote Sensing Approach," *International Journal of Remote Sensing* 31, no. 10 (2010).
[52] Alexander De Juan, "Long-Term Environmental Change and Geographical Patterns of Violence in Darfur, 2003–2005," *Political Geography* 45 (2015).
[53] Michael Kevane and Leslie Gray, "Darfur: Rainfall and Conflict," *Environmental Research Letters* 3, no. 3 (2008).

Finally, De Waal gives concludes that "depleted natural resources and livelihood transformations cannot on their own account for armed conflict." [54] Just like all other explanations, an explanation solely focusing on one cause - the resource scarcity in Darfur in this case - oversimplifies the conflict and does not do justice to the political and historical reality of the region.

Counter-insurgency

The third explanation perceives the genocide as a counter-insurgency warfare, and thus as a reaction to the attacks to the government's military bases.

A report of The International Committee of Inquiry into Darfur (ICID) concluded that "those who planned and organised attacks on villages pursued the intent to drive the victims from their homes, primarily for purposes of counter-insurgency warfare." [55] Additionally, via the ruthlessness in Darfur, Sudan's regime could discourage insurgents in the east of Sudan, where rebels already controlled a large enclave on the Eritrean frontier.[56] De Waal subscribes that conclusion and notes that since time immemorial, "the Sudan government's strategy for pursuing its counter-insurgency has set every major campaign down a particular path."[57] He exemplifies this with the counter-insurgency of 1985 when due to "the huge financial costs of mobilising the army, the political unpopularity of the draft, and the uncertain loyalty of many army officers, the government made the fateful decision to use proxy militia."[58] Five other Sudanese counter-insurgency operations have embarked on the same path. For similar reasons as in 1985, Sudan chose the same path to curb the rebellion in Darfur. Sudan's involvement in South Sudan already drove their army to the extreme, pushing financial and social limits. An additional draft for an unpopular and costly war in Darfur would face hostility from Sudan's population. Therefore, the Sudanese government partnered with the Janjaweed, creating the latest addition to the traditional counter-insurgencies.

Sceptics of the explanatory power of counter-insurgency demonstrate how the alleged counter-insurgency operation in Darfur does not fit into accepted models of counter-insurgencies. According to these general theories, counter-insurgencies are aimed selectively towards the civilian base of rebel groups.[59] However, violence in Darfur was aimed at non-Arab civilians indiscriminately. Additionally, the regimes' response was not proportionate to the rebel's attacks.

---

[54] Alex de Waal, as quoted in: Declan Butler, "Darfur's Climate Roots Challenged," *Nature* 447, no. 1038 (2007).
[55] "Report of the International Commission of Inquiry on Darfur to the United Nations Secretary-General ", (Geneva: United Nations, 2005), 4.
[56] Ryle, 62.
[57] Alex De Waal, "Reflections on the Difficulties of Defining Darfur's Crisis as Genocide," *Harvard Human Rights Journal* 20 (2007): 27.
[58] Ibid., 28.
[59] For an important overview, see: Benjamin Valentino, Paul Huth, and Dylan Balch-Lindsay, "'Draining the Sea': Mass Killing and Guerrilla Warfare," *International Organization* 58, no. 02 (2004).

These seeming inconsistencies between practice and theory are easily parried because Darfur's violence was born out of an alliance between the government and a militia. This compromise between political and ethnic goals of the Janjaweed and the counter-insurgency interests of the central government explains the 'ineffectiveness' of the violence. The characteristics of this compromise clearly resemble characteristics of previous Sudanese counter-insurgencies.[60]

Statistical studies do are not conclusive about the applicability of the theory on counter-insurgency either. Hagan critiqued the relevance of counter-insurgency by concluding that "there is no or little evidence from surveys to support … that the attacks have been directed at rebel groups as a counter-insurgency strategy."[61] In a study of a small region in Darfur, he observes that sites of rebel presence do not correlate with sites of violence.[62] This, however, does not imply that crushing the insurgents was not a prime aim of Sudan's regime. The regime has repeatedly bombed places where rebel leaders have met to hammer out differences, clearly showing that the violence was also specifically focused on crushing and disorganising the rebellion.[63] This disruption of the rebel's organisation continuously inhibited peace talks, as rebels were internally divided and thus did not have a strong negotiating position towards the Sudanese government. The observed, and more comprehensive, violence of the Janjaweed also crushed the insurgents, but with a lot of collateral damage. As Hagan fails to acknowledge, the missing correlation between rebel presence and violence can be explained via the characteristics of the alliance between the government of Sudan and the Janjaweed; for the government of Sudan, crushing the insurgents was only possible by allowing the Janjaweed agency over the execution of the violence.

[60] Shaw.
[61] Hagan, Rymond-Richmond, and Parker, 525.
[62] Hagan and Kaiser.
[63] Eric Reeves, "Genocide without End?: The Destruction of Darfur," *Dissent* 54, no. 3 (2007): 10.

Intra-elite rivalry

The fourth explanation relates closely to the arguments for counter-insurgency warfare. As van der Maat first proposed, elites often use the momentum of genocides to strengthen their own pillars of support and purge intra-elite rivals.[64] Hassan al-Turabi, the vice-president and an influential Islamist thinker, was such an intra-elite rival of Omar al-Bashir. Al-Turabi aspired to expand the Arab identity beyond the elites in Khartoum, into the rest of Sudan.[65] He had inflamed the sentiment against al-Bashir, who allegedly thought African Darfurians could never belong to the Arab ethnicity. Later, Al-Bashir accused Al-Turabi of having ties with JEM.[66] Al-Bashir explicitly blamed al-Turabi for inciting the rebellion and escalating the war in Darfur. The Sudanese dictator closed al-Turabi's party and had him imprisoned several times in the 2000s, which gradually eroded his power in Sudanese politics.[67] The International Crisis Group even argued that the power-struggle between al-Turabi and al-Bashir "for control over the Islamist movement and the country was being played out in Darfur."[68] Current research studies if al-Bashir used the Darfur genocide to purge rivalling voices from the central government and if the Darfur genocide fits into van der Maat's general theory of genocidal consolidation.

In conclusion, all presented literature has its strengths and weaknesses. Yet these explanations cannot be analysed separately as each proposed dynamic an explanation influences, or at least interfaces, with dynamics of other explanations. If taken collectively, the literature fits into a single narrative. The four theories suggest that the Darfur genocide was a counter-insurgency campaign against rebels, which simultaneously had the political purpose of ruining the reputation of political opponents of al-Bashir. This campaign took on massive dimensions due to the ethnic history of Darfur; economically, environmentally and socially marginalised tribes in the periphery fought the well-off elites in Khartoum. This intra-regional conflict affected the historical struggle between Arab and non-Arab tribes within the region as well, which has intensified over the last decades due to dwindling natural resources. As explanations are not mutually exclusive, the Darfur genocide is a typical case of overdetermination.

---

[64] For a full version of van der Maat's theory, see: Eelco van der Maat, "Genocidal Consolidation: Final Solutions to Elite Rivalry," http://dx.doi.org/10.2139/ssrn.2665673.
[65] Alex De Waal, "Who Are the Darfurians? Arab and African Identities, Violence and External Engagement," *African affairs* 104, no. 415 (2005): 191.
[66] Yehudit Ronen, "The Rise and Fall of Hasan Abdallah Al-Turabi: A Unique Chapter in Sudan's Political History (1989–99)," *Middle Eastern Studies* 50, no. 6 (2014): 1001.
[67] Ibid.
[68] "Darfur Rising: Sudan's New Crisis," in *ICG Africa Report* (Brussels: International Crisis Group, 2004), 9.

3. RESEARCH DESIGN

As the reading of literature showed, weighing and comparing the relative importance of rival explanations has proven to be problematic. This thesis aims to determine this relative importance through a statistical analysis, to answer to what extent ethnic tensions and resource scarcity explains the severity of violence in Darfur. On the basis of publicly available geo-coded data, this research design seeks to correlate the alternative explanations' observable and quantifiable implications with spatiotemporal data on episodes of violence. SPSS is the chosen software package for all statistical work. QGIS, an open-source GIS software, processes all the geographical data.[69]

Because the availability of useful geo-coded data is limited and observable implications are not easily quantifiable, only two out of the four previously presented theories can be tested: genocide as an ethnic conflict and genocide as a struggle over scarce resources. Ethnicity and resource scarcity can have interdependencies. The explanations are not mutually exclusive; as an example, ruling ethnic groups might have located themselves around the most fertile lands. Retrospective analysis of these linear interdependencies is a key aspect of result interpretation.

The chosen method brings strong advantages; a statistical study offers an objective way of comparing explanations while simultaneously avoiding a collection of qualitative examples as the circumstantial evidence for a single theory. Furthermore, using GIS and geo-coded data offers a unique combination of usable data and easy processing of large amounts of data.

One of the gravest limitations to this study is the limited availability of data, which affects the research design in several ways. Firstly, the limited availability of geo-coded data on Darfur fuels the need to approximate observable implications and to make assumptions of varying acceptability. Decisions on independent variables heavily depend on the availability of data. This limited availability often forces us to measure observable implications more deviously. Although a lot of valuable data is unavailable, the selected data contains no gaps or missing entries and thus describes the entire surface of Darfur, another unique aspect of this analysis.

The remainder of this chapter first addresses the selection of the to-be-tested theories. The statistical strategy then outlines how the statistical analysis can answer the research question. The section that follows selects the dependent and independent variables as well and identifies the selection and processing of several datatypes: baseline data, data on the dependent variable

[69] Elaborate explanations on the workings of these software packages are not presented in this thesis, as they largely depend on the version and operating system. Most explanations or manuals can be found online.

and data on independent variables. The data on the independent variables quantifies the observable implications of the tested theories. The same section hypothesises how variations the independent variables influence the dependent variables. These hypotheses are tested in the results. The chapter concludes with a short overview of the complete research design, including observable implications and the selected methods, scope and data.

Selection of theories

Two of the proposed explanations for the Darfur genocide are excluded from this research. Firstly, previous statistical studies already scrutinised the view of the genocide as a Sudanese counter-insurgency.[70] Although these studies have qualitative shortcomings, publicly available data cannot mitigate the known flaws. As the review of literature previously argued, Khartoum's aim of the counter-insurgency has been entangled and diluted with the practicality of its coalition with the Janjaweed. This entanglement complicates quantification of the observable implication of the theory and thus also determination of the importance of counter-insurgency in the Darfur genocide.

Arguably, ethnicity can be an observable implication of the counter-insurgency theory as rebel groups mainly recruited from Fur, Massaleit or Zaghawa groups. Yet, this approach cannot discriminate between the two theories on ethnicity and on counter-insurgency. Furthermore, it would be overly simplistic to denote ethnicity as a variable for counter-insurgency efforts as Hagan, Rymond-Richmond and Parker identified a variation of rebel activity between villages with similar ethnic composition.[71] Combined, these two complications make it infeasible to statistically test the importance of counter-insurgency during the Darfur genocide.

Secondly, the observable implications of intra-elite rivalries, such as attempted coups or sackings, arrests or resignation of high-ranking officers within the regime, cannot be quantified to an extent that makes data usable for this analysis; there is simply no data on the spatial variation of support for al-Bashir or his rivals within Darfur. Intra-elite rivalry therefore remains black-boxed.

An initial scan of the availability of data on observable implications showed that the remaining two theories are testable. Therefore, this study only focuses on the relevance of ethnic tensions and resource scarcity to the Darfur genocide.

[70] Hagan, Rymond-Richmond, and Parker.
[71] Ibid., 548.

Statistical strategy

The statistical strategy is paramount to successfully answering the research question. The statistical strategy essentially determines how the selected data is processed and analysed and how interpretation of the results can answer the research question.

The first issue of the statistical strategy is to define the statistical units into which all the selected data must be processed.[72] Collectively, all statistical units must cover the entire area of Darfur, fitting together like pieces of a jigsaw puzzle. The aim of data collection and processing is to quantify every observable implication into a variable in each statistical unit. The statistical analysis is then performed on the complete set of statistical units.

The number of statistical units must be sufficiently high to form a representative sample size; a necessity for obtaining precise results. The units' surface area units can thus not be too large, as the statistical analysis would then become imprecise. However, the surface area of units cannot be too small either as the data, which are often published in grid cells, then does not adequately describe variations between statistical units. Statistical units of ten square kilometres cannot adequately describe variations of data in grid cells of hundred square kilometres. In other words, the resolution of the statistical units must roughly match the resolution of the input data.

The sixty-three counties of Darfur meet these criteria and are therefore selected as the statistical units. This choice is beneficial as the number of counties allows for a good statistical evaluation and input data is often organised along similar administrative boundaries or, alternatively, roughly equally sized grid cells. All the input data must be processed into variables per county. Figure 1 shows the location of Darfur's counties within Sudan.

[72] For an overview on statistical units, see: Shaila Nijhowne, "Defining and Classifying Statistical Units," *Business Survey Methods* 214 (2011).

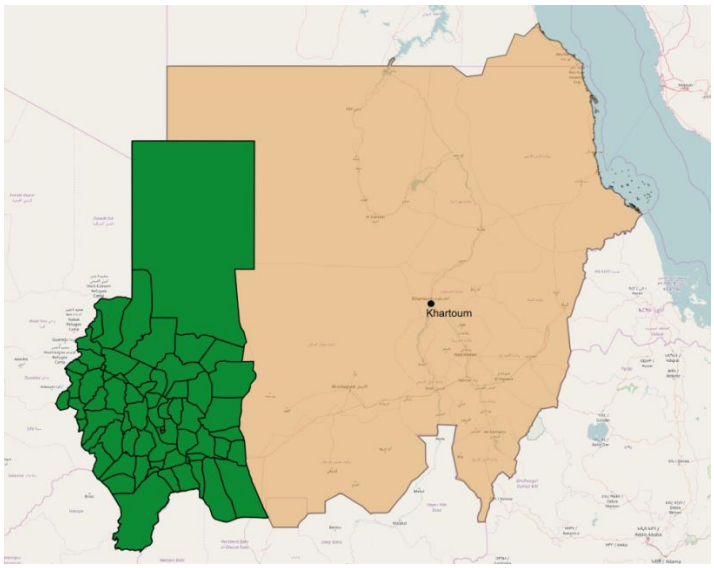

Figure 1: The location of Darfur's sixty-three counties (in green) within Sudan.

To generalise conclusions from statistical units to Darfur in its entirety, the study uses two (one per explanation) linear regression models with the least squares approach for the fit. These linear regressions should find correlations between independent and dependent variables. This chapter's conclusion gives a mathematical overview of the regression models and statistical strategy, including all the selected variables.

A linear regression model is based on two key assumptions, which are validated in hindsight. Firstly, the error between the observed and predicted values should be normally distributed. Residual plots can assess the normality of errors. These plots also show if a linear model is the right choice for the data. In the residual plot, the plotted errors between the data and the regression model should be random, not systematic. A second assumption is that linear interdependencies, or multicollinearity, within the data is limited. As an indication of interdependencies, the Variance Inflation Factor (VIF) score of each variable should be lower than three.

The two regression models thus produce results that allow for comparison of predictive strength and accuracy of two sets of variables, corresponding to the observable implications of two theories. Yet, it is impossible to directly compare two regressions because the units and dimensions of variables vary. To allow for a comparison of the two regressions despite this issue, all variables are standardised.[73] The standardised value $\tilde{x} = \frac{x-\mu}{\sigma}$, in which μ is the average value off x and σ is its standard deviation. Assuming a normal distribution of variables, the standardised values should lie between around -4 and +4 for every single variable. The

[73] For a basic overview of statistics, and the use of standardisation and Z-scores, see: Timothy C Urdan, *Statistics in Plain English* (New York: Routledge, 2016), 37 - 49.

standardisation ensures that results, such as regression coefficients, of the two different regression models are comparable. Eventually these results can demonstrate a significant difference in predictive strength between the two theories.

Interpreting the results is key to understanding the practical relevance of the regression models. Firstly, the coefficient of determination (R^2 value) indicates the average difference between data points and the regression model. The closer R^2 is to unity, the better the model describes variations within the data and the more accurate the model is. For comparison between the several regression models, the adjusted R^2 will be compared to account for the difference in the number of independent variables.

If the coefficient of determination shows that the model describes the data accurately, the p-values for regression coefficients can then assess if they are significant ($p < 0.05$) or not ($p > 0.05$). Regression coefficients are the extent to which variations of a specific independent variable change the value of the dependent variable.

Additionally, the Pearson correlation coefficients tell whether the dependent and independent variables are negatively or positively correlated, or, alternatively, not related at all. The correlation coefficient ranges from -1 to 1 and signals pure negative linearity at -1, no linearity at zero and pure positive linearity at 1. A Fisher r-to-z transformation then assigns Z-values to the difference between correlation coefficients, thereby testing the significance of the difference between the coefficients.[74] A Z-value of 1.64 signals a statistically significant difference between coefficients with a p-value of 0.05. This analysis indicates the statistical significance of the difference between explanatory power of several variables.

Still, the model itself will always remain a proof of a correlation and not necessarily of a causal relation. However, the right selection and processing of data in combination with lessons from the literature ensures that these correlations approximate the qualitative truth as accurately as possible.

[74] Xiao-Li Meng, Robert Rosenthal, and Donald B Rubin, "Comparing Correlated Correlation Coefficients," *Psychological bulletin* 111, no. 1 (1992); N Clayton Silver and William P Dunlap, "Averaging Correlation Coefficients: Should Fisher's Z Transformation Be Used?," *Journal of Applied Psychology* 72, no. 1 (1987).

Data selection

The data-selection is categorised into three phases: baseline data, data on the dependent variable and data on the independent variable. The baseline data merely processes and categorises all other data and therefore has no influence or relevance for the actual results. Data on the independent variable describes the data on the dependent variable. The created regression models show how the latter correlates to the former. This section goes through all the decisions concerning selection, collection and processing of these variables.

Baseline data

Baseline data forms a framework to process raw data into the independent variables. As the sixty-three counties form the statistical units, one type of baseline data describes the administrative boundaries of these counties as established in 2009 and confirmed by the UN Office for the Coordination of Humanitarian Affairs (see Figure 1).[75] The number of data entries is thus fixed at sixty-three, which is sufficiently high to run regression models. The surface area of each county is also calculated from the area that the counties' boundary encloses. This area averages other data out to produce variables as population density.

Data on population and population density divides out the impact that population has on results. Logically violence is more likely to occur in densely populated areas. Dividing independent variables by population can offset these effects. The used data is taken from the Gridded Population of the World dataset of the year 2000, published by NASA.[76] The dataset forms a heat map of the world, subdivided into grid cells of one square kilometre. Every pixel value of the heat map indicates the total population of that square kilometre. The dataset is clipped to the size of Darfur and then all the pixels within each county are summated to calculate the total population per county. Dividing the population by the surface area of the county gives the population density per county in Figure 2.

[75] OCHA Sudan, "Sudan - Administrative Boundaries (1 and 2)," UN OCHA, https://data.humdata.org/dataset/sudan-administrative-boundaries.
[76] Erin Doxsey-Whitfield et al., "Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4," *Papers in Applied Geography* 1, no. 3 (2015).

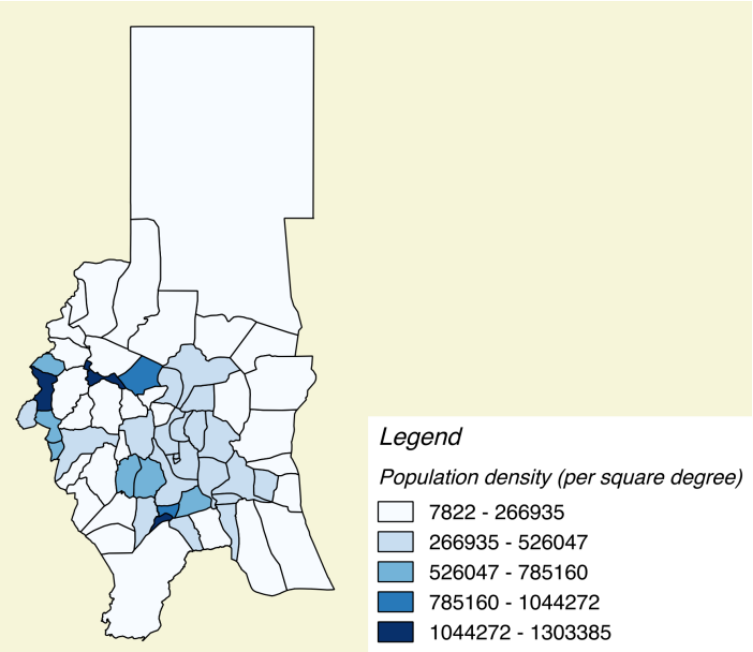

Figure 2: Population density per county in Darfur

The dependent variable

The selection of the dependent variable is a crucial element of the research design as well. This study uses unique geo-coded data on 9.183 villages in Darfur gathered by the U.S government.[77] The data collection aimed to substantiate the priorities of humanitarian agencies. The data marks each village as destroyed, damaged or intact. A village is 'destroyed' when more than half of its structures are destroyed and 'damaged' when less than half of its structures are destroyed. The data clearly shows the horrific intensity of the used violence. Eventually, in 2010, almost forty percent of all villages were destroyed or damaged. Figure 3 also shows that the violence was most severe in the first years of the conflict and its intensity reduced significantly after 2007, when the first UN peacekeeping was deployed.

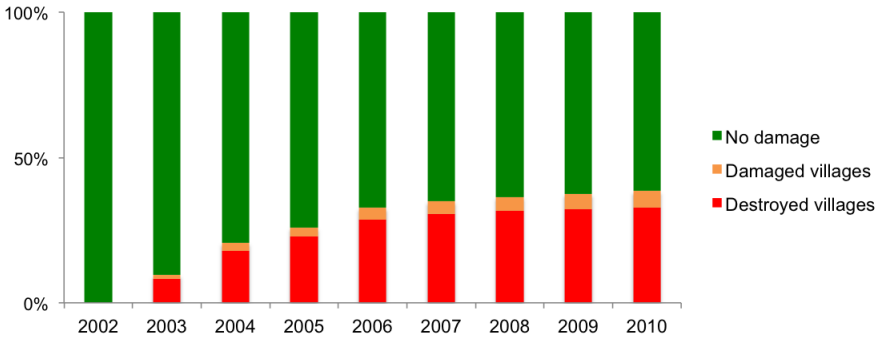

Figure 3: The status (cumulative) of all of Darfur's villages through time.

[77] Samuel Totten and Eric Markusen, "The Us Government Darfur Genocide Investigation," *Journal of Genocide Research* 7, no. 2 (2005); U.S. Department of State - Humanitarian Information Unit, "Darfur Damaged and Destroyed Villages," U.S. Government, https://data.humdata.org/dataset/darfur-damaged-and-destroyed-villages.

Alongside its geographical position and its status, the data describes the number of destroyed structures per village, the total number of structures per village and the approximate year in which of the destruction. The data incorporates all confirmed events between February 2003 and December 2010, when fresh peace talks between rebels and central Sudanese government started. This same timeframe will thus apply to the validity of the results. The conclusions thus do not necessarily apply to any dynamics prior or after this time window.

This data shows both the temporal and the spatial variation of the violence. An analysis of the temporal variation of the violence, with the year of destruction as the dependent variable, would require temporal variation within the independent variables as well. Time-dependent effects as migration fluxes or varying presence of humanitarian aid workers would have to be mapped on multiple moments during the conflict. However, a first scan of the data on all Darfur's villages shows some temporal variation, which forms an interesting starting point for future research (see left of Figure 4). An analysis of temporal variation of the violence could give insights into the effects of the deployment of a peacekeeping mission and the effect of humanitarian organisations. For now, conducting a temporal analysis is far beyond the scope of this research and would simply be setting up for failure.

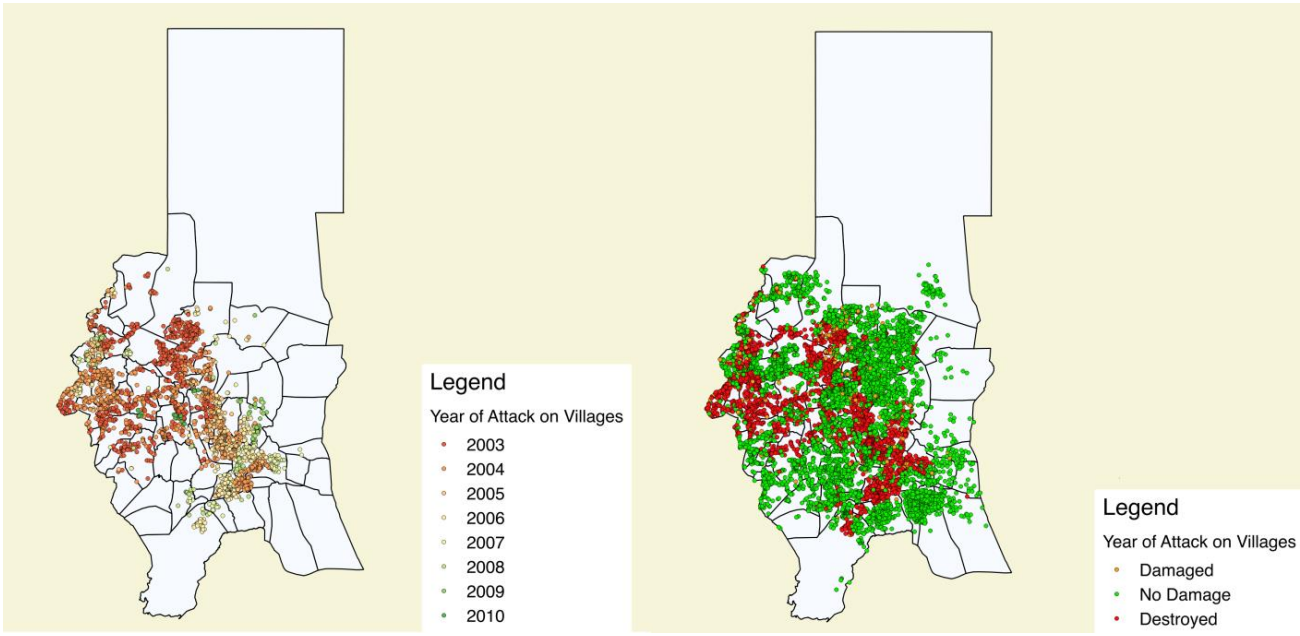

Figure 4: Left: the location of all damaged and destroyed villages coloured on the basis of the year of the attacks. **Right:** the location of all villages coloured on the basis of their status.

Therefore, this study chooses to analyse the spatial variation of the severity of violence per county, thereby excluding all temporal elements. The most logical measure for the severity of violence is the percentage of structures that were destroyed:

$$y_i = \left(\frac{\sum_{j=0}^{j=m} Number\ of\ destroyed\ structures_j}{\sum_{j=0}^{j=m} Number\ of\ structures_j} \right)_i ,$$

in which each village *j* is a part of a larger collection of *m* villages in county *i*.

This dependent variable y_i would have several strengths. Firstly, the variable accounts for the number of structures in villages that were not attacked. Secondly, this variable offsets disturbances via population sizes as counties with a large population have more structures as well. Without this offset, the population is likely to blur the correlation as logically violence concentrates more in populated than desolated areas.

Sadly, this dependent variable is not obtainable from the data in our dataset, as the number of structures per village is only known for 2.707 out of 9.183 villages, roughly thirty percent. As the villages with an unknown number of structures are mainly undamaged and heterogeneously distributed over the region, analysing only the remaining 2.707 villages would infect the whole model with a strong selection bias. To mitigate this problem, a new variable is introduced. The percentage of destroyed structures within each village *j* is now denoted as:

$$\varphi_j = \frac{Number\ of\ destroyed\ structures_j}{Number\ of\ structures_j}.$$

When no data on the number of structures is available for that specific village, the φ comes to depend on the status of the village, using the fact that every village is categorised into 'no damage', 'damaged' or 'destroyed' (see Figure 5). When a village *j* is marked as 'No Damage', φ_j equals zero. When a village *j* is 'Damaged', φ assumes the average value of φ for all damaged villages with data on structures, so that $\varphi_j = 0.407$. Completely destroyed villages get assigned $\varphi_j = 0.964$, the average φ of all 'destroyed' villages with data on structures. In this way, missing data points are filled up, without altering the distribution of the original data.

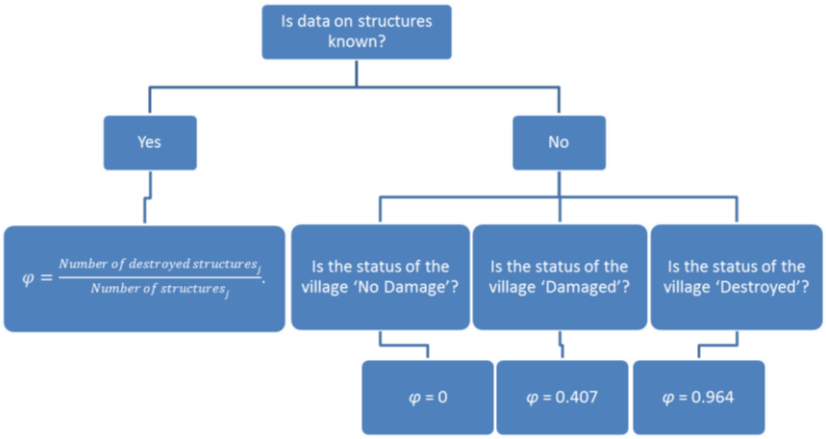

Figure 5: Decision tree for calculation of φ

As this φ_j is calculated per village, it still needs to be generalised to the county-level. To do so, the φ-values of all villages m in each county i are summated. To then account for different population sizes per county, the summated φ-value is divided by the population of county i (Pop_i) to obtain the eventual dependent variable y_i:

$$y_i = \frac{\sum_{j=0}^{j=m} \varphi_j}{Pop_i}$$

Figure 6 shows the distribution of y over Darfur. It is remarkable that nine counties have experienced no violence whatsoever. Furthermore, it seems that the violence was mostly concentrated in Central and West Darfur.

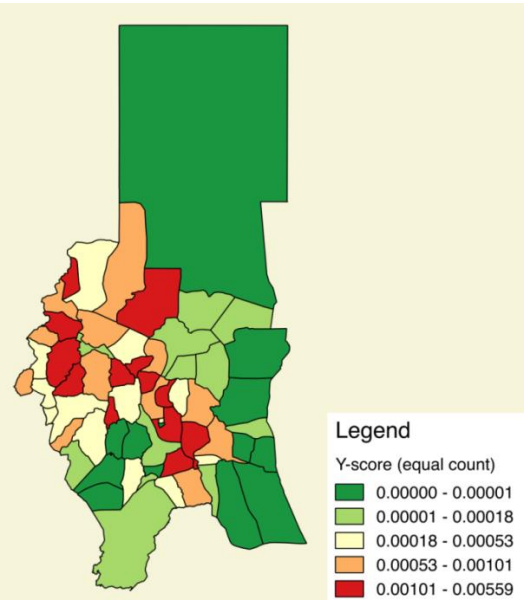

Figure 6: Distribution of y over Darfur's counties. The colour classes are divided into equal counts.

Data for alternative explanations

Now that the base-line data and the dependent variable are discussed, the focus must shift towards the independent variables. Two sets of independent variables are used; one for ethnicity and one for resource scarcity.

Ethnicity

The Fur, Massaleit and Zaghawa (FMZ) tribes constituted the main targets of governmental marginalisation over the last centuries. Before that, these non-Arab tribes were the rich, privileged and ruling tribes of Darfur. These tribes are also the ones from which the rebellion recruited their support. If ethnicity explains the severity of violence in Darfur, the population density of these three tribes would be a strong predictor of the severity of violence. The hypothesis is thus that the more densely counties are inhabited by an FMZ-population, the more severe the violence in that county is.[78]

Several data sources had to be combined to calculate $E_{FMZ,i}$, the FMZ-population density in county *i*. The GeoEPR dataset presented the boundaries of the ethnic territories of minorities.[79] This dataset with global coverage has geo-coded every ethnic group on a map. Due to the absence of dataset specifically focusing on Darfur, this global map is the state-of-the-art concerning ethnic boundaries in the region. Where other global dataset's boundaries are strictly separate, GeoEPR allows for the presence of multiple ethnicities in a single location. For the FMZ-groups, the data of the GeoEPR claims to be valid between 1956 and 2013, thus also during the time of the Darfur genocide. As this dataset excludes refugee fluxes or changes in identity, the presented ethnic boundaries are assumed to have been constant throughout the conflict.

GeoEPR thus provides polygons and surface areas in which the FMZ-population lives. The population count of each tribe comes from the Joshua Project.[80] As it is intrinsically troublesome to count how many people identify with an ethnic group, these estimates are one of the few available data sources. The FMZ-population counts are from 2014 and thus do not correspond to the counts of 2003. However, these effects are neglected, as the focus is the distribution of E_{FMZ} rather than its exact value.

[78] See the literature review and the introduction of this thesis for a full description of the role of ethnicity and, especially, the role of Fur, Massaleit and Zaghawa tribes in Darfur. Additional information can be found in: Boggero; De Waal, "Who Are the Darfurians? Arab and African Identities, Violence and External Engagement."
[79] Julian Wucherpfennig et al., "Politically Relevant Ethnic Groups across Space and Time: Introducing the Geoepr Dataset " *Conflict Management and Peace Science* 28, no. 5 (2011).
[80] "Joshua Project: Sudan," Joshua Project, https://joshuaproject.net/countries/SU.

Several steps and assumptions are made to further process the data towards an analysable variable. The calculations assume that each ethnic population is distributed homogeneously over each associated polygon. This assumption simplifies the computation of E_{FMZ} significantly as now there is a linear dependency between area and population. Although this assumption does not describe reality accurately, it allows for calculation of E_{FMZ}, which approximates the truth better than alternatives such as a dummy variable (0 or 1) for the presence of ethnic groups.

Now, the total FMZ-population per county can be calculated through a simple method. First, the total area of each FMZ-polygon is calculated. Then, per county, the area that each FMZ-polygon covers is calculated. When an FMZ-polygon completely covers a county, this area will equal the county's area. When a county has no overlap with any FMZ-polygon, this area will equal zero. The FMZ-population in county *i* of a total of *j* counties is then calculated as follows:

$$Pop_{FMZ,i} = \frac{A_{F,i}}{\sum_{i=0}^{i=j} A_{F,i}} * \sum_{i=0}^{i=j} Pop_{F,i} + \frac{A_{M,i}}{\sum_{i=0}^{i=j} A_{M,i}} * \sum_{i=0}^{i=j} Pop_{M,i} + \frac{A_{Z,i}}{\sum_{i=0}^{i=j} A_{Z,i}} * \sum_{i=0}^{i=j} Pop_{Z,i} ,$$

in which *Pop* denotes population and A is the surface area. To compute $E_{\mathrm{FMZ,i}}$, the FMZ-population in a county is simply divided by that county's surface area. Figure 7 shows the spatial variation of $E_{\mathrm{FMZ,i}}$ over Darfur. The dimensions of $E_{\mathrm{FMZ,i}}$ do not merely approximate actual numbers as the total population of Fur, Massaleit or Zaghawa is very hard to measure and as several assumptions were necessary to arrive at these figures. Furthermore, E_{FMZ} cannot distinguish between violence against Fur, Massaleit or Zaghawa groups, as it is a summation over these three groups. Despite these limitations, Figure 7 gives a good representation of the relative distribution of the FMZ-population over Darfur's counties.

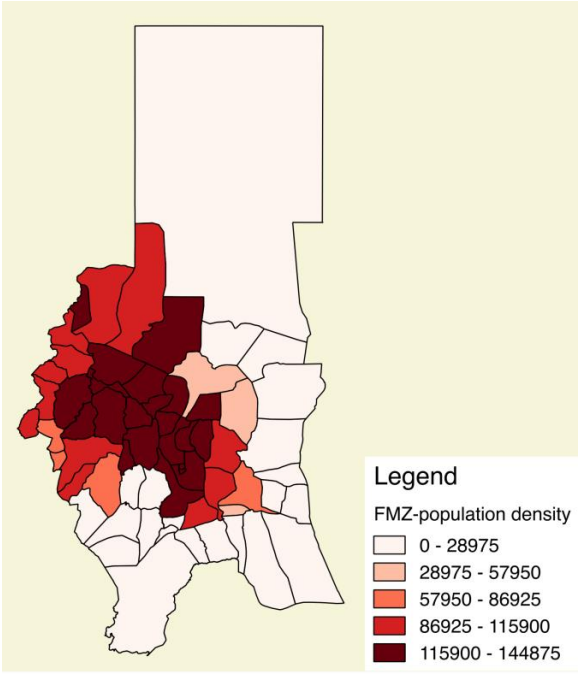

Figure 7: FMZ-population density per county in Darfur, measured in FMZ-population per square degree of surface

As an additional measure of ethnic tensions, the minimal distance from each county to the centroid of the closest ethnic polygon (Fur, Massaleit or Zaghawa) is calculated. This D_{FMZ} gives an indication of the proximity of each county to possible ethnic tensions. The hypothesis is that the severity of violence increases when a county is closer to an ethnic centroid. Because we are focusing on positive predictors to *y*, Figure 8 shows the inverted distance instead of the regular distance.

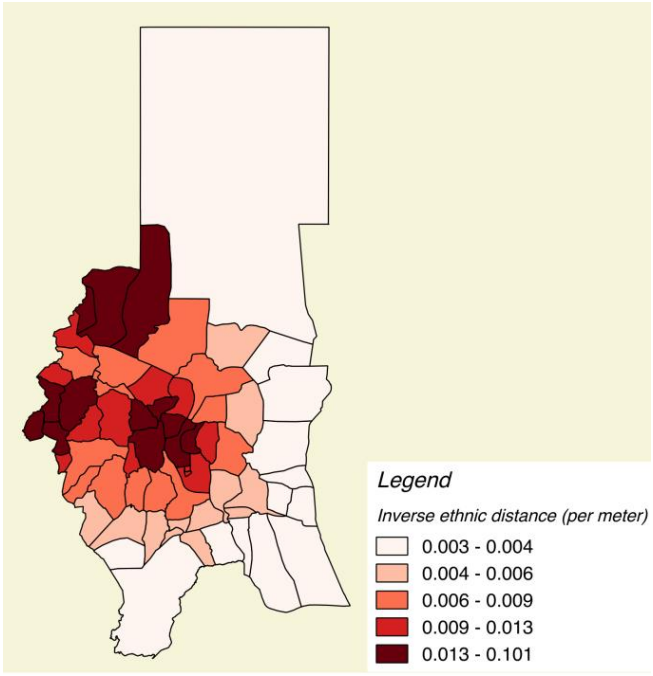

Figure 8: The distance between counties and the centres of areas with FMZ-population.

Marginalisation along ethnic lines, as an integral part of the theory on ethnicity, is excluded from this analysis for several reasons. Firstly, the effect of marginalisation on the severity of violence is very indirect. Including variables for marginalisation would essentially hypothesise that these variables can predict ethnic divisions, which then predicts the severity of the violence. Yet, other reasons for the marginalisation of specific counties in Darfur other than just ethnicity might exist; ecological, historical or institutional factors can also significantly influence observable implications of marginalisation as infant mortality, literacy rates, malnutrition or poverty.

Secondly, although variables of marginalisation indirectly describe the severity of violence, its variables could also make opposite hypotheses very plausible. Take, as an example, the access to infrastructure, quantified as the kilometres of road per square kilometre per inhabitant. For this variable, two opposite hypotheses are both plausible. Access to infrastructure can affect the severity of violence negatively, assuming that counties with worse access to infrastructure are more marginalised, which increases the likelihood that people from FMZ-tribes inhabit the county and thus the county would experience severer violence; a very indirect and unwieldy hypothesis. An opposite effect of access to infrastructure is also plausible: the better people's access to infrastructure in a county, the better the militia's access to the people, the severer the violence. There is thus no clear and direct hypothesis that ties observable implications of marginalisation to the severity of violence.

Thirdly, data on marginalisation is scarce and not available in the right resolutions. Data on inequality, poverty, malnutrition, literacy rates, gender equality or access to education, health care or judicial support is only available as a single figure for Darfur, thus not differentiating for different counties. Because of these three reasons, the explanatory power of ethnicity is only tested on the basis of the E_{FMZ} and D_{FMZ}.

Resource scarcity

The literature suggests that resource scarcity is another main contributor to Darfur's violence. The regression model for this theory comprises two sets of variables. The first set of independent variables describes the resource scarcity as each county's population experiences it.

Again, limited availability of data heavily influences the decisions on independent variables to represent the observable implications of this theory. Variables for land capture or theft of livestock would describe the motivation of militiamen most directly, but such data is unavailable. Furthermore, data on food scarcity is not available in a sufficient resolution and data on population growth is only available for years after the conflict. Therefore, this research

quantifies the availability of water, the fertility of soil and the severity of droughts for people in each county as independent variables. The hypothesis around these variables is always: if there is more competition around resources in a county, then the violence in that county will be severer.

The data used to describe water scarcity comes from the Aqueduct Global Maps of the World Resource Institute.[81] This data is organised along boundaries that relate to hydrological areas as basins, rivers, deserts or *wadis*, delivering a sufficiently high resolution to average out this data towards the county-level. Amongst others, this data maps the Baseline Water Stress (*BWS*), which measures the total annual water withdrawal as a percentage of the total amount of available blue water (see left of Figure 9). A high *BWS* indicates high competition amongst users and would correlate with a high severity of violence according to our hypothesis.

Furthermore, the right of Figure 9 presents an index for the severity of droughts. This Drought Severity Index (*DSI*) is calculated from a variable *M* in the raw data, which is the mean length of droughts multiplied with the severity of droughts. This variable *M* is then corrected for the county's total population and surface area, to measure the competition amongst the local population, so that the *DSI* is high when competition over water is high and many people experience extreme drought in a small patch of land:

$$DSI_{\mathrm{i}} = \frac{M_{\mathrm{i}} * Pop_{\mathrm{i}}}{A_{\mathrm{i}}}$$

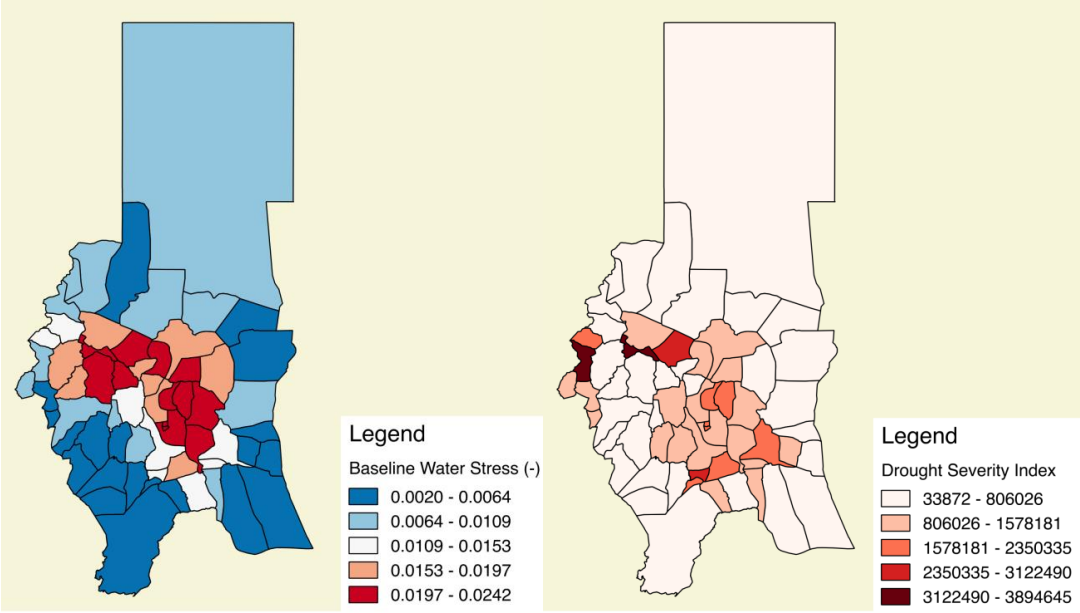

Figure 9: *BWS* in all counties of Darfur (left) and the *DSI* (right).

[81] Francis Gassert et al., "Aqueduct Global Maps 2.1: Constructing Decision-Relevant Global Water Risk Indicators," in *Working Paper* (Washington, DC: World Resources Institute, 2014).

Soil fertility is the main determinant for the opportunity to cultivate land for agriculture, thus also for the value and potential of land. The associated hypothesis claims that counties with fertile soil face more competition over land and will thus have severer violence.

The soil fertility is measured via data on vegetation. Several studies suggest a causal relation between the two; if the vegetation is denser, then the soil is more fertile as well.[82] The data comes from the Normalised Difference Vegetation Index, as mapped by the NASA Earth Observatory.[83] The data shows the vegetation on grid cells of one hundred square kilometres for February 2003; the month in which the documentation of the severity of violence started. The data comes in a high resolution and is first averaged over the surface area of each county. Then, the population density of each county is divided by its vegetation index to express the competition over the fertile soil in a new index. This new Soil Competition Index (*SCI*) is thus high when many people have to share a small patch of barren soil. Figure 10 plots the *SCI* per county.

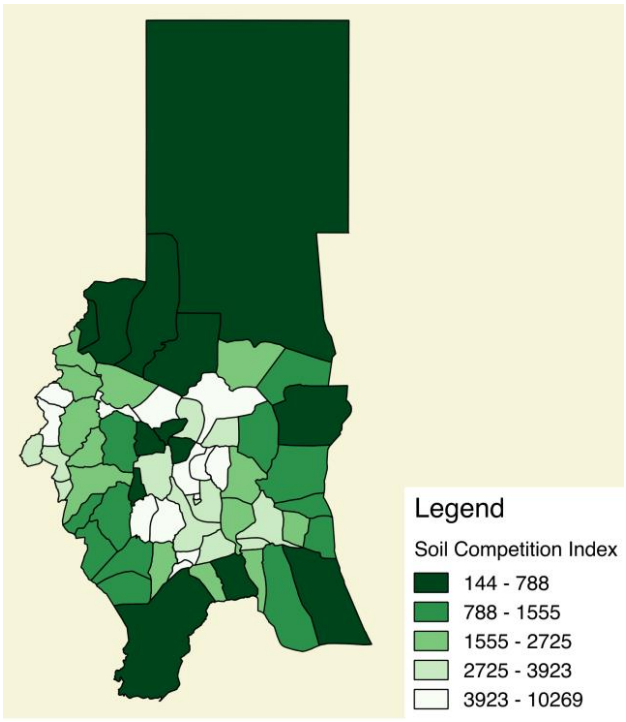

Figure 10: Soil Competition Index. The whiter counties show counties where the competition over fertile soil is fiercer.

[82] Patricia Guidão Cruz Ruggiero et al., "Soil-Vegetation Relationships in Cerrado (Brazilian Savanna) and Semideciduous Forest, Southeastern Brazil," *Plant Ecology* 160, no. 1 (2002); D. D. Eni, A. I. Iwara, and R. A. Offiong, "Analysis of Soil-Vegetation Interrelationships in a South-Southern Secondary Forest of Nigeria," *International Journal of Forestry Research* 2012 (2012).
[83] John Weier and David Herring, "Measuring Vegetation (Ndvi & Evi)," NASA's Earth Observatory Group, https://earthobservatory.nasa.gov/Features/MeasuringVegetation/.

The second set of variables describes the role of oil per county. This set indicates the importance of each county's resources to the central government, rather than to the local population. Two indicators measure the importance of oil per county, both originating from United States Agency for International Development.[84] The presence of oil production sites is a dummy variable, with each zero indicating an absence and each one indicating an oil production site within that county ($N_{oilpsite}$).

The second variable for the relevance of oil is calculated through the oil concession zones. Oil concession zones are areas of which the government obtained the rights to explore and produce oil. To process these areas into useful data, all polygons are first combined into one. Then, the percentage of each county's surface that overlaps with oil concession areas is calculated (P_{oil}). Figure 11 summarises the distribution of P_{oil} in Darfur, a good indication of each county's relevance in the central government's plans to explore and produce oil.

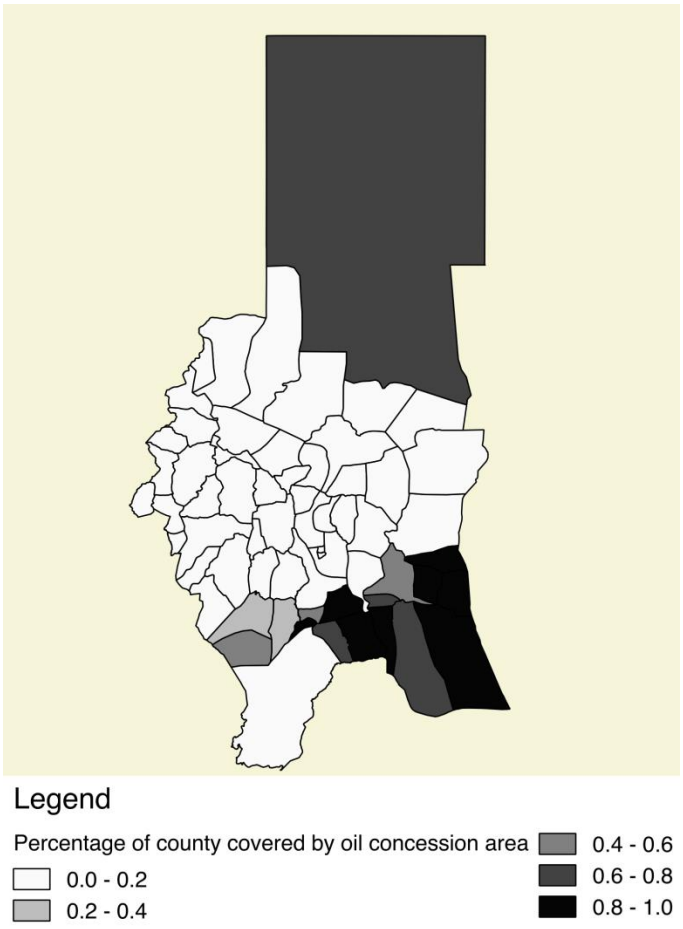

Figure 11: The percentage of each county that is covered by oil concession areas (P_{oil}), for whole of Darfur.

[84] Geosprocket LLC, "An Assessment of Existing Spatial Data in South Sudan," USAID Global Environment Management Support Program, https://github.com/GeoSprocket/ssudan-eco.

To conclude, we expect the severity of the violence (y) to be a function of variables that describe the observable implications of two alternative theories: ethnicity and resource scarcity. Both sets of variables are tested in two separate regression models. The first model aims to test if violence resulted from ethnic tensions and accounts for FMZ-population density (E_{FMZ}) and the inverse distance between counties and the closest ethnic centre (D_{FMZ}). In further use, this model is presented as Model 1:

$$y_i = \beta_{0,a} + \beta_{1,a} E_{FMZ,\text{i}} + \beta_{2,a} D_{FMZ,\text{i}} + \varepsilon.$$

The second model aims to tie the severity of violence to resource scarcity, with the Baseline Water Stress (*BWS*), the Drought Severity Index (*DSI*), the competition over fertile soil (*SCI*), the coverage by oil concession areas (P_{oil}) and the number of oil production sites (N_{oilpsite}). In further use, this model is presented as Model 2:

$$y_i = \beta_{0,b} + \beta_{1,b} BWS_{\text{i}} + \beta_{2,b} DRO_{\text{i}} + \beta_{3,b} SCI_{\text{i}} + \beta_{4,b} P_{\text{oil,i}} + \beta_{5,b} N_{\text{oil,i}} + \varepsilon.$$

In both models, ε is a small number indicating small random errors that deviate from the linear relation; the β's are the regression coefficients of which β_0 accounts for any violence when all other variables assume a value of zero. The models thus suggest a linear relation between the regression coefficients and the dependent variables, in which all the independent variables contribute to the severity of violence. The comparison of results for both regression models can answer the research question and thus show if the Darfur genocide was primarily a campaign of ethnic cleansing, a struggle over scarce resources or both.

4. RESULTS AND DISCUSSION

The following section presents and compares the results of the two regression models. All presented results apply to the standardised variables. Although many additional conclusions and observations can be made on basis of the presented data, this thesis only discusses contributions to the main research question.

Model accuracy

Table 1 shows the coefficients of determination for both models. The coefficients show a relatively weak fit of the regression model to the dependent variable for Model 1. The value of R^2 shows that merely 32.9% of the variance in *y* is predictable from variance in the independent variables. Yet R^2 values above 50% are rare in sciences aiming to describe human behaviour.

Table 1: The model accuracy of Model 1 and Model 2

	R	***R*²**	**Adjusted *R*²**	**Standard Error of Estimation**
Model 1	0.574	0.329	0.307	0.833
Model 2	0.550	0.302	0.241	0.871

Model 2 describes variances in the severity of violence even less adequately. To compare two models with a different number of input variables, the adjusted R^2-value must be evaluated. The adjusted R^2-values show that Model 1 can account for more of the variance in the severity of violence than Model 2 does. Therefore, variables for ethnic tensions seem to have more explanatory power towards the severity of violence than variables for resource scarcity do, although the gap is minimal. This gap closes further when the variables that account for the presence of oil are discarded (see Appendix A). This shows that the variables of oil harm the accuracy and precision of Model 2. The presence of oil does not seem to be a good predictor of the severity of violence in Darfur.

Residual plots and normality checks

Normal P-P plots check normality assumption, which dictates that errors around the expected values should be distributed normally. Figure 12 shows no major reason for concern.

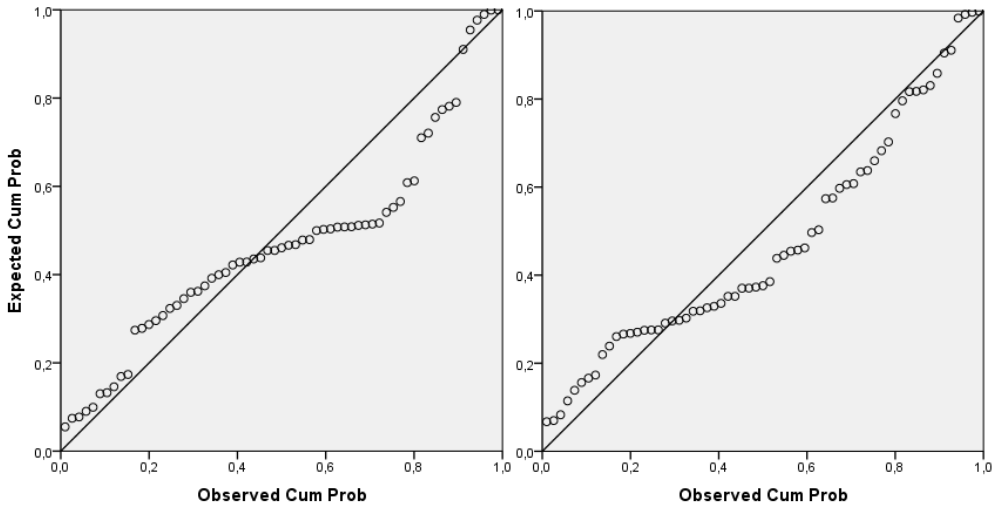

Figure 12: Normal P-P plot of Regression Standardised Residual for Model 1 (left) and Model 2 (right)

Residual plots form an additional check on the normality assumption and if a linear model is appropriate. The points should form a random pattern around the line $y = 0$, and especially between $-0.75 < x < 0.75$. This seems to be the case although there are more positive than negative outliers. The plots are slightly imbalanced around the *y*-axis, although not worryingly much. Both models show large amounts of heteroscedasticity, implying that residuals get larger as the standardised predicted value increases. The models are thus more 'off' as the severity of violence increases. This is not an inherent problem, but merely a sign that the model leaves room for improvements through normalising the errors via appropriate variable transformation. However, these kinds of transformations are beyond the scope of this research and, additionally, severely complicate the interpretation of the regression's results.[85] The plots do not show non-linearity. Overall, these two figures show that a linear model is appropriate and roughly conforms to the normality assumption.

[85] A seemingly appropriate transformation would be the Box-Cox transformation: Jason W Osborne, "Improving Your Data Transformations: Applying the Box-Cox Transformation," *Practical Assessment, Research & Evaluation* 15, no. 12 (2010).

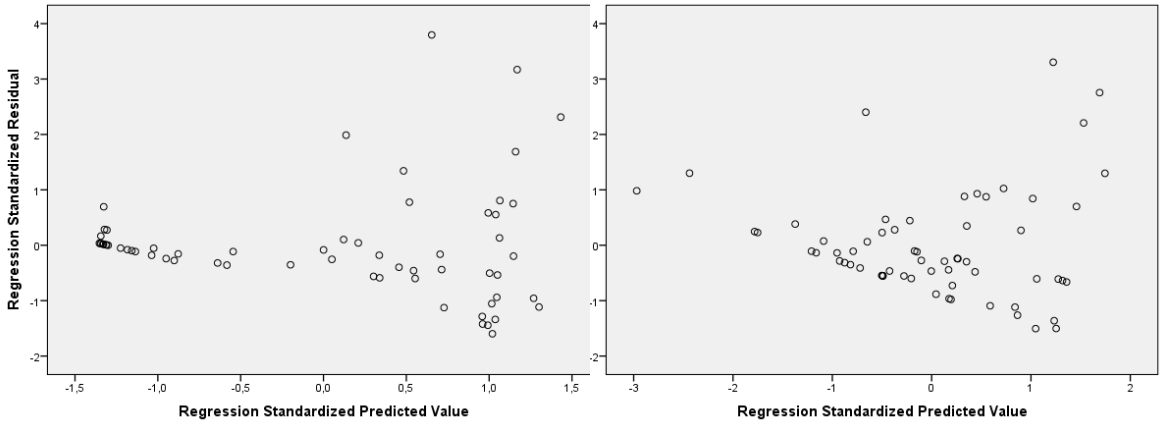

Figure 13: Standard residual plots of Model 1 (left) and Model 2 (right)

Pearson correlations coefficients

The Pearson correlations coefficients show the linearity of the correlation between different variables. If the coefficients is more extreme (-1 or 1), a linear relation captures the data more accurately. This, however, does not ensure that this linearity can be quantified in a regression coefficient, as we will see later. Because all variables are standardised, the Pearson coefficient is equal to the regression coefficient and to R in the scenario that the regression model only consists of the corresponding independent variable and y. Therefore, discrepancies between the Pearson coefficient and the regression coefficients show how other independent variables in the model impact the correlation of a specific variable to y. Comparing the Pearson coefficient thus gives a good indicator of the strength of individual correlations.

As expected, an increase of ethnic variables leads to a significant increase in the severity of violence. Table 2 shows that the coefficient between y and E_{FMZ} is by far the largest of all, meaning that the impact of E_{FMZ} on the severity of violence is the most linear and the most positive. The coefficient between y and D_{FMZ} is less positive.

Coefficients for Model 2 show several unexpected results. Firstly, it was expected that resource scarcity would increase the severity of violence, but the Pearson coefficients do not subscribe this hypothesis unambiguously. Some coefficients are negative while they were expected to be positive contributors to the severity of violence. *SCI*, for instance, has a negative correlation to the severity of violence. This suggests, in contrast to our hypothesis, that when more people have to share a smaller patch of more barren soil, the violence will become less severe. An alternative variable, which increases as more people share a smaller patch of *fertile* soil instead of *arid* soil, was tested as well but did not significantly alter the results (see Appendix B). An alternative hypothesis thus might be more applicable to understand the relation between soil competition and severity of violence.

Several other coefficients are insignificant and thus no conclusion can be drawn from them. In addition, Table 2 shows that there is no relation between the severity of violence and the number of oil sites, as the coefficient of determination for Model 2 without oil variables had already suggested. In conclusion, the results from Model 2 are ambiguous and do not conform to earlier hypotheses. The exception is *BWS*, whose coefficients are both positive and significant.

Table 2: Pearson correlation coefficient between all variables, with ***y*** in **bold** and insignificant coefficients ($p > 0.005$) in red

	y	$E_{FMZ,i}$	$D_{FMZ,i}$			
y	-	**0.569**	**0.388**			
E_{FMZ}	**0.569**	-	0.584			
D_{FMZ}	**0.199**	0.584	-			
	y	***BWS***	***DSI***	***SCI***	P_{oil}	$N_{oilpsite}$
y	-	**0.342**	**-0.027**	**-0.219**	**-0.280**	**-0.126**
BWS	**0.342**	-	-0.162	0.371	-0.368	-0.224
DSI	**-0.027**	-0.162	-	0.021	0.024	0.356
SCI	**-0.219**	0.371	0.021	-	-0.099	-0.149
P_{oil}	**-0.280**	-0.368	0.024	-0.099	-	0.173
$N_{oilpsite}$	**-0.126**	-0.224	0.356	-0.149	0.173	-

The coefficient of determination already showed that Model 1 describes the data more accurately than Model 2 does. In addition to that, the Pearson coefficients thus show that E_{FMZ} of Model 1 is the strongest positive contributor to the severity of violence, while the coefficients of Model 2 show ambiguity and disarray. Further evaluation must show if the difference between the coefficients of E_{FMZ} and other variables is indeed significant.

Regression Coefficients

The Pearson coefficients had already shown that both variables in Model 1 are positive contributors to the severity of violence. The Pearson coefficients also showed that E_{FMZ} has a stronger linear relation to y than D_{FM}. Table 3 confirms these two findings; the regression coefficient for E_{FMZ} is again the largest of all while the regression coefficient for D_{FMZ} is not significant at all. Yet, a separate analysis on the change in R^2 when the variables of Model 1 are added to the model one by one also shows that both variables are significant contributors to the model's fit (see Appendix C).

Table 3: Regression coefficients for Model 1 and Model 2

Model 1	β	Standard Error for β	t	p-score	95% confidence intervals for β		Collinearity statistics	
					Lower bound	Upper Bound	Tolerance	VIF
Constant	0	0.110	0.000	1.000				
E_{FMZ}	0.521	0.130	3.996	0.000	0.260	0.781	0.659	1.518
D_{FMZ}	0.084	0.130	0.084	0.524	-0.177	0.344	0.659	1.518
Model 2	β	Standard Error for β	t	p-score	95% confidence intervals for β		Collinearity statistics	
					Lower bound	Upper Bound	Tolerance	VIF
Constant	0	0.110	0.000	1.000				
BWS	0.444	0.130	3.410	0.001	0.183	0.705	0.721	1.387
DSI	0.153	0.120	1.276	0.207	-0.087	0.394	0.848	1.180
SCI	-0.419	0.121	-3.475	0.001	-0.660	-0.177	0.842	1.187
P_{oil}	-0.141	0.120	-1.173	0.246	-0.382	0.100	0.848	1.180
$N_{oilpsite}$	-0.119	0.122	-0.974	0.334	-0.363	0.126	0.821	1.217

For several reasons, the regression coefficients for Model 2 are again ambiguous. First, three of the regression coefficients are not significant (*DSI*, P_{oil}, $N_{oilpsite}$), of which two also had an insignificant Pearson coefficient. From these results conclude that droughts and the presence of oil do not have a measurable impact on the severity of violence.

The only significant contributors to Model 2 are *BWS* and *SCI*. Opposed to our hypothesis but conform its Pearson coefficient, *SCI* correlates negatively to *y*. For *BWS* the correlation to *y* is positive as expected. A separate analysis, which records the change that occurs to R^2 when the variables of Model 2 are added to the model one at a time, also shows that *BWS* and *SCI* are the only significant contributors to the model (see Appendix D).

The Variance Inflation Factor (VIF) in Table 3 measures the multicollinearity between the variables. As the VIFs for all variables are lower than three, thereby these models conform to one of the main assumptions of a linear regression, namely that variables must be linearly independent.

Discussion

On the basis of these results, the impacts of several variables on the severity of violence are deemed insignificant or non-existing (D_{FMZ}, *DSI*, P_{oil} and $N_{oilpsite}$). The results for *SCI* do not agree with prior hypotheses. As this discrepancy erodes the qualitative value of any conclusion, this variable is excluded from further analyses as well. Table 4 summarises the remaining variables, both positive and significant contributors to their respective models.

Table 4: Comparison of significant regression coefficients

	Pearson coefficient (with *y*)	β	Standard Error for β	t	p-score	95% confidence intervals for β		Collinearity statistics	
						Lower bound	Upper Bound	Tolerance	VIF
E_{FMZ}	0.569	0.521	0.130	3.996	0.000	0.260	0.781	0.659	1.518
BWS	0.342	0.514	0.139	3.702	0.000	0.236	0.791	0.622	1.608

To compare the predictive power of these two variables more directly, both variables are added stepwise to a new regression model, Model 3. The coefficient of determination is recorded for E_{FMZ} solely first and then again for both variables after the addition of *BWS* to the new model. The same procedure is repeated with opposite sequence of addition of the variables (first *BWS* and then E_{FMZ}).

Table 5 shows that for the first sequence, the addition of *BWS* does not contribute to the model's accuracy, as the change in R^2 is insignificant. In the second sequence, despite *BWS*'s advantage of being added first, the addition of E_{FMZ} improves the model a lot. Comparing the R^2-change of 3^c and 3^d shows that the addition of E_{FMZ} improves the model more than the addition of *BWS* to an empty model. This shows that E_{FMZ} totally dominates the explanatory power of this combined model.

Table 5a: Summaries of the coefficients of determination for different models.

Model	R	R^2	Adjusted R^2	R^2 change	F change	Sig F change
3[a]	0.569	0.324	0.313	0.324	29.278	0.000
3[b]	0.574	0.329	0.307	0.005	0.432	0.513
3[c]	0.342	0.117	0.103	0.117	8.098	0.006
3[d]	0.574	0.329	0.307	0.212	18.956	0.000

a. Predictors: Constant, E_{FMZ}

b. Predictors: Constant, E_{FMZ}, BWS

c. Predictors: Constant, BWS

d. Predictors: Constant, BWS, E_{FMZ}

Table 6 presents the regression coefficients for this new model. These coefficients confirm the dominance of E_{FMZ} in the new model as the regression coefficient for BWS is insignificant. Yet again, the regression coefficient for E_{FMZ} is the sole contributor to the explanatory power of the new regression model. Also note that the Pearson coefficients with y have not changed compared to early models as we are dealing with standardised variables.

Table 6: Summary of key coefficients of Model 3

	Pearson coefficient (with y)	Standard Error (for Pearson)	β	Standard Error for β	t	p-score for β	95% confidence intervals for β		Collinearity statistics
							Lower bound	Upper Bound	VIF
E_{FMZ}	0.569	0.105	0.636	0.146	4.354	0.000	0.344	0.928	1.906
BWS	0.342	0.120	-0.096	0.146	-0.657	0.513	-0.388	0.196	1.906

A final comparison of the two Pearson coefficients for E_{FMZ} and *BWS* (r_{FMZ} and r_{BWS}) can add further significance to the dominant role of E_{FMZ} in predicting the severity of violence. The hypothesis for this comparison is that $r_{FMZ} > r_{BWS}$. The hypothesis is accepted with a one-tailed significance level of 0.05 when the Z-score for the comparison of the $Z_{comparison} > 1.64$. A Fischer r-to-z-transformation allows for a comparison of r_{FMZ} and r_{BWS}.[86] The transformation is relatively easy to perform:

$$z = \frac{1}{2} * \ln\left(\frac{1+r}{1-r}\right)$$

The two Z-scores that are obtained from this equation are then filled into the following equation, which computes $Z_{comparison}$:

$$Z_{\text{comparison}} = \frac{(z_1 - z_2)}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}},$$

with N_1 and N_2 being the sample sizes, i.e. the number of counties. This Z-score assesses the significance of the difference between two Pearson coefficients. For our coefficients, $Z_{comparison} = 1.59$, just below the critical value of 1.64. As $Z_{comparison} < 1.64$, our hypothesis that $r_{FMZ} > r_{BWS}$ with a one-tailed significance level of 0.05 must be rejected. Instead, the hypothesis $r_{FMZ} > r_{BWS}$ has a significance level of 0.056.

[86] Clayton Silver and William Dunlap, "Averaging Correlation Coefficients: Should Fisher's R-to-Z Transformation Be Used?," *Journal of Applied Psychology* 72, no. 1 (1987).

5. CONCLUSIONS

The most incredible number regarding this research is that the war in Darfur has devastated the lives of millions of people. Scholarly interest, arising from these atrocities, has suggested four theories that explain the causes of the violence in Darfur. As these theories are not mutually exclusive, weighing and comparing them has been problematic. On the basis of geo-coded data, this thesis offers a statistical comparison between two of the explanations for the violence in Darfur: ethnicity and resource scarcity. Simply put, the main research question is: which of two explanations, ethnicity or resource scarcity, contributed more to the root causes of the violence in Darfur?

The research design statistically ties observable implications of the selected theories to the severity of violence in Darfur. In this statistical analysis, the independent variables describe these observable implications and the dependent variable is the severity of violence. Data on these variables is collected and calculated for each of the sixty-three counties of Darfur, which are the statistical units of this research. Seven independent variables and a single dependent variable are correlated in two linear regression models, one for each explanation. The accuracy, strength and precision of the two models can then reveal identify the main contributor to the violence in Darfur; ethnicity, resource scarcity, both or neither.

This research design has many benefits as it is objective and allows for a unique combination of publicly available, geo-coded data. Yet, it has obvious limitations as well. Firstly, the literature showed that both selected theories have gathered widespread academic support. The literature also exposed the main limitation of the research, namely that the two untested explanations are equally plausible. The limited availability of data compelled the exclusion of the other two theories, but a more comprehensive study should also include variables on counter-insurgency and intra-elite rivalry.

The low availability of data severely limits this research in other ways as well. The data scarcity has forced choices for variables that indirectly describe the theories' observable implications, as data for a more direct description was unavailable. In addition, the data scarcity necessitates assumptions in the processing and calculation of variables.

Despite its benefits and limitations, the research design partly answers the research question. The evidence is not demonstrative, but it is probable that the model for ethnicity (Model 1) better predicts the severity of violence than the model for resource scarcity (Model 2) does. Several statistical arguments underpin this conclusion. Firstly, according to the coefficients of determination, Model 1 better describes variances in the severity of violence better than Model 2. Secondly, the Pearson coefficients for Model 1 are positive and significant, while the same coefficients for Model 2 are either negative or insignificant, with the exception of *BWS*. This tells that variables for Model 1 have a stronger positive linearity to the severity of violence than variables for Model 2.

The regression coefficients tell a similar story. The regression coefficient of E_{FMZ} is the largest of all variables, but the regression coefficient of D_{FMZ} is insignificant. Regression coefficients of Model 2 show a similar trend as the Pearson coefficients of Model 2; regression coefficients are either negative or insignificant, with the exception of *BWS*. While the insignificant coefficients identify variables that do not impact the severity of violence, the negative coefficients demonstrate inconsistencies with prior hypotheses. The prior hypotheses expected that when resources are scarcer, violence would be severer. Instead, some coefficients suggest that violence is less severe when resources are scarcer. The literature revealed similar discrepancies; some studies proved a correlation between resource scarcity and violence while others dismissed those conclusions.[87] The discrepancies with the hypotheses and the insignificance other coefficients put serious doubts on the explanatory power of Model 2 as a whole. Apparently, variables of Model 2 cannot be quantitatively or qualitatively tied to the severity of violence.

The only positive and significant contributor to Model 2 is the *BWS*. So overall, only E_{FMZ} and *BWS* have significant and positive Pearson coefficients and regression coefficients. To directly compare their predictive power to the severity of violence, both variables are combined into a single new model: Model 3. *BWS*'s contribution to Model 3 falls into insignificance when compared to E_{FMZ}'s contribution. Model 3 thus relies heavily on E_{FMZ} for its accuracy and fit.

[87] See page 10: Brown. De Juan. Kevane and Gray; Butler.

The Pearson coefficient for E_{FMZ} (r_{FMZ}) seemed larger than Pearson coefficients for *BWS* (r_{BWS}) and all other variables. However, the Fisher r-to-z-transformation and the succeeding comparison of the Z-scores, show that the difference between r_{FMZ} and r_{BWS} is only significant with a *p*-score of 0.056, just above the pre-determined significance level of 0.05. The hypothesis $r_{FMZ} > r_{BWS}$ is thus rejected. Therefore, the conclusion that ethnicity explains more of the severity of violence than resource scarcity lacks a closing piece of statistical evidence.

Finally, the previously formulated hypotheses are reassessed. Regarding those hypotheses, this statistical analysis concludes that the severity of violence (i) indeed increases with the population density of Fur, Massaleit and Zaghawa tribes, (ii) might increase with the counties' proximity to ethnic centres, (iii) is not significantly impacted by resource scarcity (with exception of *BWS*) and (iv) is not significantly impacted by the presence of oil.

Overall, ethnicity is probably, but not definitely, a stronger predictor of the severity of violence than resource scarcity. The coefficients of determination, the Pearson coefficients, the regression coefficient, the discrepancies and inconsistencies of Model 2, the results for Model 3 and the reassessment of the primary hypotheses all point into that direction. Although all statistical analyses point towards ethnicity - and especially E_{FMZ} - as the strongest predictor, the final evidence is inconclusive as the difference between Pearson coefficients for E_{FMZ} and *BWS* is statistically insignificant.

Apart from this main conclusion, the research elucidates interesting individual correlations. According to this research, the presence of oil and the severity of droughts have been irrelevant to this war. The significantly negative correlation between *SCI* and the severity of violence shows that violence is more severe in areas where a few people share a lot of fertile land. These correlations could form a starting point for new qualitative research, focusing on the dynamics of violence in fertile, but desolated areas of Darfur.

In retrospect, this research also unveils the complexity of the genocide in Darfur. It was very challenging to capture the dynamics of the genocide into variables, partly due to the limited availability of data. Despite many efforts, not a lot of data has a proper resolution *and* is publicly available *and* is trustworthy *and* is geo-coded. Therefore, on-the-ground surveys and population counts would benefit analyses of the Darfur genocide. Although many scholars have conducted surveys and counts, none of them has disclosed their data in usable formats.[88] Research with a longer timespan can aim to set up a data-sharing platform, to optimally benefit from all available

[88] Hagan and Kaiser; Hagan and Rymond-Richmond, "The Collective Dynamics of Racial Dehumanization and Genocidal Victimization in Darfur."; *Darfur and the Crime of Genocide*; Jennifer Leaning MD.

knowledge. Furthermore, an increased timespan might also allow researchers to access data that is not publicly available. As the dust on the Darfur genocide settles, more data will become available on the dynamics of the violence. Variables on land and cattle capture, changes of ethnic composition of counties or changes in the distribution of key positions in government could then shine new light on this genocide.

Until now, the full potential of data in humanitarian crises is yet to be unravelled. But in the future, as data continues to becomes more abundant, available and determinative, these kinds of data-driven analyses might prove to be of paramount importance. Data-driven decision-making can ensure an adequate and fast response from the international community and can optimise the allocation of humanitarian organisations' aid. In addition, it can serve as evidence for genocides and help to hold its killers accountable so that dictators like Omar Al-Bashir will not go unpunished and so that, in the future, genocides as Darfur's will be a thing of the past.

APPENDIX A

Comparison of Model 2 and Model 2 without variables for oil presence

	R	R^2	Adjusted R^2	Standard Error of Estimation
Model 1	0.574	0.329	**0.307**	0.833
Model 2	0.506	0.256	**0.190**	0.900
Model 2 – without oil	0.520	0.270	**0.233**	0.876

	β	Standard Error for β	p-score	95% confidence intervals for β		Collinearity statistics	
				Lower bound	Upper Bound	Tolerance	VIF
BWS	0.515	0.122	0.001	0.271	0.759	0.833	1.200
DSI	0.119	0.526	0.297	-0.107	0.345	0.966	1.035
SCI	-0.413	0.120	0.001	-0.653	-0.172	0.855	1.169

APPENDIX B

Comparison of Model 2 with the old *SCI* and the new *SCI*

Model 2	β	Standard Error for β	t	p-score	95% confidence intervals for β		Collinearity statistics	
					Lower bound	Upper Bound	Tolerance	VIF
BWS	0.444	0.130	3.410	0.001	0.183	0.705	0.721	1.387
DSI	0.153	0.120	1.276	0.207	-0.087	0.394	0.848	1.180
SCI	-0.419	0.121	-3.475	0.001	-0.660	-0.177	0.842	1.187
P_{oil}	-0.141	0.120	-1.173	0.246	-0.382	0.100	0.848	1.180
$N_{oilpsite}$	-0.119	0.122	-0.974	0.334	-0.363	0.126	0.821	1.217

Model 2b	β	Standard Error for β	t	p-score	95% confidence intervals for β		Collinearity statistics	
					Lower bound	Upper Bound	Tolerance	VIF
BWS	0.434	0.137	3.166	0.002	0.160	0.709	0.675	1.481
DSI	-0.311	0.137	-2.271	0.027	-0.586	-0.037	0.675	1.482
SCI-B	**-3.94**	**0.129**	**-3.059**	**0.003**	**-0.652**	**-0.136**	**0.765**	**1.308**
P_{oil}	-0.027	0.132	-0.206	0.838	-0.291	0.237	0.729	1.327
$N_{oilpsite}$	-0.063	0.117	-0.536	0.594	-0.296	0.171	0.931	1.074

APPENDIX C

Overview of contribution of both variables of Model 1 to the overall fit of the model.

Model	R	R^2	Adjusted R^2	R^2 change	F change	Sig F change
1	,388[a]	,150	,136	,150	10,794	,002
2	,574[b]	,329	,307	,179	15,965	,000

a. Predictors: Constant, D_{FMZ}

b. Predictors: Constant, D_{FMZ}, E_{FMZ}

c. Dependent Variable: *y*

APPENDIX D

Overview of contribution of all variables of Model 2 to the overall fit of the model.

Model	R	R^2	Adjusted R^2	R^2 change	F change	Sig F change
1	,342[a]	,117	,103	,117	8,098	,006
2	,375[b]	,141	,112	,024	1,644	,205
3	,554[c]	,306	,271	,166	14,098	,000
4	,560[d]	,314	,267	,008	,645	,425
5	,563[e]	,317	,258	,003	,282	,598

a. Predictors: Constant, *BWS*

b. Predictors: Constant, *BWS*, *DSI*

c. Predictors: Constant, *BWS*, *DSI*, *SCI*

d. Predictors: (Constant), *BWS*, *DSI*, *SCI*, P_{oil}

e. Predictors: (Constant), *BWS*, *DSI*, *SCI*, P_{oil}, $N_{oilpsite}$

f. Dependent Variable: *y*

BIBLIOGRAPHY

Abdul-Jalil, Musa Adam. "The Dynamics of Customary Land Tenure and Natural Resource Management in Darfur." *Land Reform FAO*, no. 2 (2006): 9 - 23.

Behrends, Andrea. "Fighting for Oil When There Is No Oil Yet: The Darfur-Chad Border." *European Journal of Anthropology* 2008, no. 52 (2008): 39-56.

Boggero, Marco. "Darfur and Chad: A Fragmented Ethnic Mosaic." *Journal of Contemporary African Studies* 27, no. 1 (2009): 21-35.

Brown, Ian. "Assessing Eco-Scarcity as a Cause of the Outbreak of Conflict in Darfur: A Remote Sensing Approach." *International Journal of Remote Sensing* 31, no. 10 (2010): 2513-20.

Bungeroth, Paul Jeffrey and Annie. "In Pictures: Darfur Refugees Then and Now." BBC, http://www.bbc.com/news/world-africa-28739000.

Butler, Declan. "Darfur's Climate Roots Challenged." *Nature* 447, no. 1038 (2007).

Cobham, Alex. "Causes of Conflict in Sudan: Testing the Black Book." *The European Journal of Development Research* 17, no. 3 (2005): 462-80.

Cruz Ruggiero, Patricia Guidão, Marco Antônio Batalha, Vânia Regina Pivello, and Sérgio Tadeu Meirelles. "Soil-Vegetation Relationships in Cerrado (Brazilian Savanna) and Semideciduous Forest, Southeastern Brazil." *Plant Ecology* 160, no. 1 (2002): 1-16.

Daly, Martin W. *Darfur's Sorrow: A History of Destruction and Genocide*. New York: Cambridge University Press, 2007.

"Darfur Rising: Sudan's New Crisis." In *ICG Africa Report*. Brussels: International Crisis Group, 2004.

De Juan, Alexander. "Long-Term Environmental Change and Geographical Patterns of Violence in Darfur, 2003–2005." *Political Geography* 45 (2015): 22-33.

De Waal, Alex. "Counter-Insurgency on the Cheap." *Review of African Political Economy* 31, no. 102 (2004): 716-25.

———. "Deep Down in Darfur: Nothing Is as We Are Told in Sudan's Killing Fields." Review of Darfur: the Ambigious Genocide by Gerard Prunier. *Review of African Political Economy* 32, no. 106 (2005): 653 - 59.

———. *Famine That Kills: Darfur, Sudan*. New York: Oxford University Press 1989.

———. "Reflections on the Difficulties of Defining Darfur's Crisis as Genocide." *Harvard Human Rights Journal* 20 (2007): 25 - 33.

———. "Who Are the Darfurians? Arab and African Identities, Violence and External Engagement." *African affairs* 104, no. 415 (2005): 181-205.

Degomme, Olivier, and Debarati Guha-Sapir. "Patterns of Mortality Rates in Darfur Conflict." *The Lancet* 375 (2010): 294-300.

Doxsey-Whitfield, Erin, Kytt MacManus, Susana B Adamo, Linda Pistolesi, John Squires, Olena Borkovska, and Sandra R Baptista. "Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4." *Papers in Applied Geography* 1, no. 3 (2015): 226-34.

Eni, D. D., A. I. Iwara, and R. A. Offiong. "Analysis of Soil-Vegetation Interrelationships in a South-Southern Secondary Forest of Nigeria." *International Journal of Forestry Research* 2012 (2012): 8.

"Entrenching Impunity: Government Responsibility for International Crimes in Darfur." Human Rights Watch, 2005.

Flint, Julie, and Alex De Waal. *Darfur: A New History of a Long War*. African Arguments. Edited by Julie Flint and Alex De Waal London: Zed Books, 2008.

Gassert, Francis, Matt Luck, Matt Landis, Paul Reig, and Tien Shiao. "Aqueduct Global Maps 2.1: Constructing Decision-Relevant Global Water Risk Indicators." In *Working Paper* Washington, DC: World Resources Institute, 2014.

Gifkins, Jess. "Darfur." In *The Oxford Handbook of the Responsibility to Protect*, edited by Alex Bellamy and Tim Dunne. Oxford: Oxford University Press, 2016.

Hagan, John, and Joshua Kaiser. "The Displaced and Dispossessed of Darfur: Explaining the Sources of a Continuing State- Led Genocide." *The British Journal of Sociology* 62, no. 1 (2011): 1-25.

Hagan, John, Wemona Rymond-Richmond, and Patricia Parker. "The Criminology of Genocide: The Death and Rape of Darfur." *Criminology* 43, no. 3 (2005): 525-62.

Hagan, John, and Wenona Rymond-Richmond. "The Collective Dynamics of Racial Dehumanization and Genocidal Victimization in Darfur." *American Sociological Review* 73, no. 6 (2008): 875-902.

———. *Darfur and the Crime of Genocide*. New York: Cambridge University Press, 2008.

Jennifer Leaning MD, SMH. "Diagnosing Genocide-the Case of Darfur." *The New England journal of medicine* 351, no. 8 (2004): 735.

"Joshua Project: Sudan." Joshua Project, https://joshuaproject.net/countries/SU.

Kaplan, Jeffrey. "A Review Of: Gerard Prunier. Darfur: The Ambigious Genocide"." *Terrorism and Political Violence* 20, no. 146 - 148 (2007).

Kevane, Michael, and Leslie Gray. "Darfur: Rainfall and Conflict." *Environmental Research Letters* 3, no. 3 (2008).

LLC, Geosprocket. "An Assessment of Existing Spatial Data in South Sudan." USAID Global Environment Management Support Program, https://github.com/GeoSprocket/ssudan-eco.

Luban, David. "Calling Genocide by Its Rightful Name: Lemkin's Word, Darfur, and the Un Report." *Chicago Journal of International Law* 7, no. 1 (2006): 303 - 20.

Medani, Khalid Mustafa. Review of Darfur's Sorrow: A History of Destruction and Genocide (New York: Cambridge University Press, 2007). *International Journal of Middle East Studies* 42 no. 1 (2010): 174-76.

Meng, Xiao-Li, Robert Rosenthal, and Donald B Rubin. "Comparing Correlated Correlation Coefficients." *Psychological bulletin* 111, no. 1 (1992): 172.

Nijhowne, Shaila. "Defining and Classifying Statistical Units." *Business Survey Methods* 214 (2011): 49.

Olsson, Ola, and Eyerusalem Siba. "Ethnic Cleansing or Resource Struggle in Darfur? An Empirical Analysis." *Journal of Development Economics* 103 (2013): 299-312.

Olsson, Ola, and Michele Valsecchi. "Quantifying Ethnic Cleansing: An Application to Darfur." (2010).

Osborne, Jason W. "Improving Your Data Transformations: Applying the Box-Cox Transformation." *Practical Assessment, Research & Evaluation* 15, no. 12 (2010): 1-9.

Prunier, Gerard. *Darfur: The Ambiguous Genocide*. Ithaca: Cornell University Press, 2005.

Reeves, Eric. "Genocide without End?: The Destruction of Darfur." *Dissent* 54, no. 3 (2007): 8-13.

———. "Getting Darfur Wrong." *Dissent* 56, no. 4 (2009): 108-12.

"Report of the International Commission of Inquiry on Darfur to the United Nations Secretary-General ". Geneva: United Nations, 2005.

Ronen, Yehudit. "The Rise and Fall of Hasan Abdallah Al-Turabi: A Unique Chapter in Sudan's Political History (1989–99)." *Middle Eastern Studies* 50, no. 6 (2014): 992-1005.

Ryle, John. "Disaster in Darfur." *New York Review of Books* 51, no. 13 (2004): 55-59.

Sachs, Jeffrey D. "Ecology and Political Upheaval." *Scientific American* 295, no. 1 (2006): 37-37.

Sharkey, Heather J. "Arab Identity and Ideology in Sudan: The Politics of Language, Ethnicity, and Race." *African Affairs* 107, no. 426 (2008): 21-43.

Shaw, Martin. "Darfur: Counter- Insurgency, Forced Displacement and Genocide." *The British Journal of Sociology* 62, no. 1 (2011): 56-61.

Silver, Clayton, and William Dunlap. "Averaging Correlation Coefficients: Should Fisher's R-to-Z Transformation Be Used?". *Journal of Applied Psychology* 72, no. 1 (1987): 146.

Silver, N Clayton, and William P Dunlap. "Averaging Correlation Coefficients: Should Fisher's Z Transformation Be Used?". *Journal of Applied Psychology* 72, no. 1 (1987): 146.

Straus, Scott. "Darfur and the Genocide Debate." *Foreign Affairs* 84, no. 1 (2005): 123 - 33.

Sudan, OCHA. "Sudan - Administrative Boundaries (1 and 2)." UN OCHA, https://data.humdata.org/dataset/sudan-administrative-boundaries.

"Sudan Post-Conflict Environmental Assessment ". 358. Nairobi: United Nations Environment Programme, 2007.

Thayer, Willa. "Sudan Bombs Rebel Areas in Darfur after Au Peacekeeper Visit." the Independent, http://www.independent.co.uk/news/world/africa/sudan-bombs-rebel-areas-in-darfur-after-au-peacekeeper-visit-430459.html.

Totten, Samuel, and Eric Markusen. "The Us Government Darfur Genocide Investigation." *Journal of Genocide Research* 7, no. 2 (2005): 279-90.

Unit, U.S. Department of State - Humanitarian Information. "Darfur Damaged and Destroyed Villages." U.S. Government, https://data.humdata.org/dataset/darfur-damaged-and-destroyed-villages.

Urdan, Timothy C. *Statistics in Plain English.* New York: Routledge, 2016.

Valentino, Benjamin, Paul Huth, and Dylan Balch-Lindsay. "“Draining the Sea”: Mass Killing and Guerrilla Warfare." *International Organization* 58, no. 02 (2004): 375-407.

van der Maat, Eelco. "Genocidal Consolidation: Final Solutions to Elite Rivalry." http://dx.doi.org/10.2139/ssrn.2665673.

Vanrooyen, Michael, Jennifer Leaning, Kirsten Johnson, Karen Hirschfeld, David Tuller, Adam C Levine, and John Hefferman. "Employment of a Livelihoods Analysis to Define

Genocide in the Darfur Region of Sudan." *Journal of Genocide Research* 10, no. 3 (2008): 343-58.

Vehnämäki, Mika. "Darfur Scorched: Looming Genocide in Western Sudan ". *Journal of Genocide Research* 8, no. 1 (2006): 51-82.

Weier, John, and David Herring. "Measuring Vegetation (Ndvi & Evi)." NASA's Earth Observatory Group, https://earthobservatory.nasa.gov/Features/MeasuringVegetation/.

Wucherpfennig, Julian, Nils B Weidmann, Luc Girardin, Lars-Erik Cederman, and Andreas Wimmer. "Politically Relevant Ethnic Groups across Space and Time: Introducing the Geoepr Dataset ". *Conflict Management and Peace Science* 28, no. 5 (2011): 423-37.