MASTERARBEIT | MASTER'S THESIS

Titel | Title

Determining Drivers of War: the Influence of Social Proximity on Militarized Interstate-Conflict

verfasst von | submitted by

Simon von Zahn B.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 913
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Applied Economics
Betreut von | Supervisor:	Assoz. Prof. Daniel Garcia PhD

"*If the mothers in every land could teach their children to understand the homes and hopes of children in every other land – in America, in Europe, in the Near East, in Asia – the cause of peace in the world would indeed be nobly served*"

– Dwight D. Eisenhower, 1954

Abstract

The contemporary global landscape is marked by displacement of people, persistent warfare, and escalating interstate conflicts. Therefore, understanding the dynamics underlying hostile behaviour between nations is of paramount importance. This thesis aims to address this pressing concern by adapting sociology's contact theory to a macro-level analysis to falsify Dwight D. Eisenhower's statement in 1954. Does contact have an influence on interstate conflict behaviour?

Contact theory, originally proposed to elucidate interpersonal relations, posits four conditions under which contact diminishes hostility: Equal status, common goals, support of authorities, and forms of intergroup cooperation (Allport, 1954). This research adopts this theory to interstate dynamics. It does so by replacing the term 'contact', which does not fit to interstate interactions, with proximity. This study defines four layers of proximity between states: social, historical, cultural, and geographical. These layers shall serve as a framework to understand the influence of proximity on international relations.

The analyses are based on one 2-Stage-Least-Squares-model (2SLS) and one Shift-Share-model (SSA). Both models rely on open-source data from various sources, such as the Stockholm International Peace Research Institute, the Uppsala University, META and the Centre for Prospective Studies and International Information (CEPII). The models investigate the impact of four layers of proximity on interstate hostility. Specifically, they consider the impact of geographical distance between conflicting parties, historical and present colonial relationships and shared languages. The analyses reveal a clear picture on the influence of proximity on interstate conflict, identifying the exact amount of variation in conflict that is explained by our proximity layers.

The findings of this study offer valuable insights into the complex drivers of interstate conflict, underscoring the need for further research to comprehensively understand and address this phenomenon. By shedding light on the underlying mechanisms shaping hostile behaviour between nations, this thesis contributes to the scholarly discourse on conflict research and offers potential avenues for development economics and peacebuilding theory on a global scale.

Zusammenfassung

Die gegenwärtige globale Landschaft ist geprägt von Vertreibung, anhaltenden Kriegen und eskalierenden zwischenstaatlichen Konflikten. Daher ist das Verständnis der Dynamik, die feindlichem Verhalten zwischen Nationen zugrunde liegt, von größter Bedeutung. In dieser Arbeit soll dieses dringende Problem angegangen werden, indem die Kontakttheorie der Soziologie auf eine Makroebene übertragen wird, um die Aussage von Dwight D. Eisenhower aus dem Jahr 1954 zu falsifizieren. Hat zwischenmenschlicher Kontakt einen Einfluss auf das zwischenstaatliche Konfliktverhalten?

Die Kontakttheorie, die ursprünglich vorgeschlagen wurde, um zwischenmenschliche Beziehungen zu erhellen, geht von vier Bedingungen aus, unter denen Kontakt Feindseligkeit abbaut: Gleicher Status, gemeinsame Ziele, Unterstützung durch Autoritäten und Formen der Zusammenarbeit zwischen Gruppen (Allport, 1954). Die vorliegende Untersuchung überträgt diese Theorie auf die zwischenstaatliche Dynamik. Dabei wird der Begriff "Kontakt", der sich nicht zur Anwendung auf zwischenstaatliche Interaktionen eignet, durch "Nähe" ersetzt. In dieser Studie werden vier Ebenen der Nähe zwischen Staaten definiert: Soziale, historische, kulturelle und geografische. Diese Ebenen sollen als Rahmen dienen, um den Einfluss von Nähe auf internationale Beziehungen zu verstehen. Die Analysen basieren auf einem 2-Stage-Least-Squares-Modell (2SLS) und einem Shift-Share-Modell (SSA). Beide Modelle stützen sich auf frei zugängliche Daten verschiedener Quellen, wie dem Stockholm International Peace Research Institute, der Universität Uppsala, META und dem Centre for Prospective Studies and International Information (CEPII). Die Modelle untersuchen die Auswirkungen von vier Ebenen der Nähe auf zwischenstaatliche Feindseligkeit. Sie berücksichtigen die Auswirkungen sozialer Vernetzung, der geografischen Entfernung zwischen den Konfliktparteien, potenziellen kolonialen Beziehungen, sowie linguistischer Nähe. Die Analysen identifizieren unsere Definition der Nähe-Ebenen als Triebkräfte zwischenstaatlicher Konflikteskalation.

Die Ergebnisse dieser Studie bieten wertvolle Einblicke in die Dynamik zwischenstaatlicher Konflikte und unterstreichen die Notwendigkeit weiterer Forschung, um dieses Phänomen umfassend zu verstehe. Diese Arbeit trägt damit zum wissenschaftlichen Diskurs in der Konfliktforschung bei und bietet potenzielle Wege für die Entwicklungsökonomie und die Theorie der Friedenskonsolidierung im globalen Maßstab.

Table of Contents

1 Introduction

The 1940s brought unprecedented destruction to European nations. 20 years after the establishment of the largest supranational institution to date, the League of Nations was not able to prevent the greatest war in history (THE CONVENANT OF THE LEAGUE OF NATIONS, 1919). In the wake of this, numerous researchers have sought to gain insight into the factors that drive conflict behaviour. Both on a micro- and a macro-level. The macro-level analysis delves into examination on the influence of different proximity layers. Caselli, Morelli, and Rohner (2015) explore how natural resources like oil can increase the likelihood of interstate conflicts. Their findings show that countries with border-adjacent oil fields are on average 3.5 times more likely to engage in bilateral conflicts compared to those without such geographical-economic conditions. Skaperdas and Syropoulos (2001) investigate how trade openness and economic interdependencies influence militarized conflict, noting that nations are more likely to arm and escalate conflicts when global market prices for goods (like butter) exceed their national prices. Their analysis suggests that economic interconnectedness tends to reduce these conflicts, as price disparities that might encourage arming disappear in more open economies. Martin, Mayer, & Thoenig (2008) challenge the assertion that trade networks directly affect conflict likelihood, arguing that while bilateral trade can reduce tensions, multilateral openness dilutes these effects by diminishing the importance of individual trade relationships. Their research indicates that increased multilateral connections lower the costs associated with reduced bilateral trade, thereby weakening the economic incentives for conflict resolution. Hess and Orphanides (2001) explore the impact of political regimes on interstate conflict, specifically testing if democracies are less prone to conflict than other regimes and finding that democratic leaders often engage in conflict due to reelection incentives. Contrary to the notion that a world of democracies would be more peaceful, they conclude that while democratic governance itself does not reduce conflict likelihood, the adherence to international norms plays a crucial role in fostering peace. While these authors have contributed to the field of conflict analysis on a macro level, they do not cover a social proximity layer. Such layer could prove sustainable in macro level conflict, providing insights on micro founded drivers of macro-level conflict. Such analysis could serve as equivalent to the research environment of intergroup conflict behaviour, to which numerous researchers have contributed to. An influential work on this regard is Gordon Allport's contact theory (1954). In his seminal work "The Nature of Prejudice", Allport outlines four essential conditions for reducing hostility in intergroup relations: equal status, common goals, intergroup cooperation, and support from authorities. He argues that equal status among groups is crucial, noting that historical efforts regarding ethnic emanzipation in the U.S. often failed to achieve true structural equality, leading to sustained prejudice and discrimination. Common goals, such as those exemplified in integrated sports teams, create a shared identity that helps diminish intergroup conflicts by aligning group interests. Moreover, Allport emphasizes that effective intergroup cooperation requires interdependent activities that foster personal connections and empathy, while the support from authorities helps legitimize and reinforce these positive interactions through legal and cultural endorsement. Approaching interstate conflict behaviour with micro founded proximity indicators might shed a light on significant drivers of large-scale conflicts and therefore also contribute to the field of international relations. Acknowledging the complexity of the subject, this analysis does not aim to cover a holistic set of potential influences on interstate behaviour. Rather, it aims to delve deeper into the social proximity approach by focussing on micro-relationships. The study also implements the cultural proximity approach by including data on common languages between states. Geographical proximity requires data on the distance between states and serves as a natural constraint on conflicts due to the fact that the cost of military operations increase over the distance at which they take place in relation to their home countries, as Martin (2023) finds. Finally, the study considers historical proximity to be an important factor to account for. This study contributes to the research environment of interstate conflict by investigating whether Allport's (1954) or Eisenhower's (1954) understanding of the influence of micro relationships on intergroup hostilities hold in bilateral interstate context. We find that the social, cultural, and geographical proximity layer, as defined in this study, do cover significant influences on interstate conflict behaviour. The results show a strong positive relationship between proximity and conflict escalation. They thus show that the calming effect of social contact on hostility, as explained by Allport (1954), does not continue at the macro level. This study therefore supplements ongoing research efforts on the influence of proximity and connectedness layers as introduced earlier (e.g. Skaperdas and Syropoulos (2001); Caselli, Morelli, and Rohner (2015)) by emphasizing on the influence of micro relationships on interstate conflict, adding a social proximity layer to the research environment.

2 Theoretical Background

Allport's understanding of intergroup social behaviour is primarily based on four defined conditions, as pointed out before. With regard to this framework, several authors added to contact theory, such as Thomas F. Pettigrew (1998). Pettigrew agreed with the basic premises of Allport's conditions but noted that they were not sufficient to cover all aspects of intergroup contact. He emphasized the importance of additional factors such as the potential for friendship and prolonged contact, which he believed were crucial for deeper transformations in attitudes. Pettigrew also introduced the idea that several psychological emotional processes occur during intergroup contact that can lead to reduced prejudice. These processes include learning about the outgroup, changing behaviour through positive interaction, generating affective ties (emotions and empathy), and ingroup reappraisal, where contact leads individuals to reassess their stereotypes and attitudes towards their own and other groups. Moreover, Pettigrew critisized Allport's theory for not specifying how the effects of contact go beyond immediate interactions. He suggested mechanisms such as decategorization, which involve reducing the salience of group identities; salient categorization, which enhances group identity in a positive context; and recategorization, which forms a new inclusive group identity. This study examines the parallels between conflict research at the micro level and the macro level. In doing so, it identifies various studies that have already investigated the influence of various proximity indicators on the conflict behavior of macro-individuals, hence states. Therefore, adapting the research framework of Allport (1954) and Pettigrew (1998).

Caselli, Morelli & Rohner (2015) analyze the impact of natural resources on interstate conflict probability. The authors build a model that uses the geographical location of oil fields in relation to borders as predictor for bilateral conflict. They therefore combine geographical proximity with economical interests that promise potential earnings and therefore trigger militarized engagements. The model reveals bilateral conflict probabilities being on avergae 3,5 times larger for country-pairs, where one country's oil fields are located directly at the border than for country-pairs with no oil fields. On the other hand, if the oil fields are located as far away as possible from the border, the difference in bilateral conflict probability fades (Caselli, Morelli, & Rohner, 2015). The combination of geographic proximity and economic incentives powerfully influence interstate conflict behavior and confirm this study's inclusion of the geographic layer. Skaperdas & Syropoulos (2001) contribute to the field with their paper ‘Guns, Butter, and Openness: On the Relationship Between Security and Trade’ that focuses on the impact of trade and economic interdependencies on militarized conflict. A central aspect of their research includes the cost of arming for open and autarch national economies. In their model, these economies are inclined to arm themselves (and escalate conflict) as soon as world prices for butter exceed their national autarkic price for butter. If again economies are open and connected, these incentives vanish due to missing price differences. Therefore, interdependencies and connectedness of national markets may indeed diminish international conflicts (Skaperdas & Syropoulos, 2001).

Martin, Mayer, & Thoenig (2008) disagree with Skapperdas & Syropoulos (2001). They argue that trade-openness does not systematically influence trends in bilateral conflict probabilities. While bilateral trade relations do tend to calm interstate conflict escalation between the states, this effect changes as soon as countries are multilaterally open and connected and the weight of single bilateral relations decreases. In this situation, the cost of reduced trading among states are lowered and therefore incentives to conflict resolution decrease as well (Martin, Mayer, & Thoenig, 2008). Trade relations and their impact on interstate conflict escalation is subject to ongoing research efforts. However, understanding trade relations as indicator for economic proximity between states represents a fertile field of research.

Hess and Orphanes (2001) concentrate on political aspects of interstate conflict. Their research framework aims to test whether democracies tend to engage in interstate militarized conflict less often than other political regimes. They find that democratic leaders tend to engage in interstate militarized conflict due to incentives caused by popularity mechanisms in reelections. They reject the hypothesis that a world consisting of democracies would be more peaceful. While democratic regimes per se do not tend to reduce conflict probability, the implementation of international norms does represent a significant driver of interstate peace (Hess & Orphanides, 2001)

3 Empirical Strategy

3.1 Data

In this section, this study will provide an overview on data sources and the respective composition of the data. First of all, getting in touch with the drivers of international conflict requires careful consideration of indicators that can measure such events. The choice fell on the fifth edition of the *Militarized Interstate Disputes* dataframe, published as part of the *Uppsala Conflict Data Program* by the *Departement of Peace and Conflict Research* at the *Uppsala Universitet* in Sweden. The Uppsala Conflict Data Program (UCDP) is the primary source of data on organized violence worldwide and the longest-running project for collecting data on civil wars, with a history of nearly 40 years. Its definition of armed conflict has become the global standard for systematically defining and studying conflicts (Department of Peace and Conflict Research, 2023). The Militarized Interstate Disputes 5.0 Dataframe (MIDB) covers data on global armed conflicts between 1814 and 2014 including multistate conflicts, interstate conflicts as well as conflicts in which only one party involved is recognized as a national state. The MIDB dataframe introduces an escalation parameter of international conflicts called HiAct (Highest action by state in dispute). It defines 23 different levels of conflict, reaching from *No militarized action* (0), as lowest escalation level, to *Join interstate war* (21), as highest possible escalation. If no information on the intensity of the conflict is available, -9 is displayed (Palmer, et al., 2020). Within the scope of our analysis, we only include armed conflicts between 1989 and 2014 in which two states engaged in conflict, and the variable *HiAct* does not show -9. *HiAct* will serve as the dependent variable in our analysis.

Measuring international micro relationships has turned out to be challenging. Early thoughts on an index based on flight traffic between countries have been discarded as well as indicators on internet traffic between countries due to representation issues. In the end, I chose the *Social Connectedness Index*, which is part of *Meta*'s *Data for Good* project. *Meta* published the Index in 2020, providing three different snapshots of user connections (as measured in facebook friendships) between 1) US States, 2) EU regions and 3) national states. For the purpose of our research question, I have included the third dataframe, covering international user connections between national states. This index measures the social connectedness between national states on a dyadic scheme, averaged over 99% of facebook users. However, there are several limitations: Countries with less than 50.000 facebook users are not included, as well as several other countries, where Facebook is banned, such as Afghanistan, Western Sahara, China, Cuba, Iraq, Israel, Iran, North Korea, Russia, Syria, Somalia, South Sudan, Sudan, Venezuela, Yemen, Crimea, Yammu and Kashmir, Donetsk, Luhansk, Sevastopol, West Bank and Gaza (Bailey, Strobel, Wong, Cao, & Kuchler, 2021). Thus, these countries are excluded from our analysis. The Social Connectedness Index is denoted by the following equation,

$$(1)\ \frac{FB\_Connections_{i,j}}{FB\_Users_i * FB\_Users_j} = Social\ Connectedness\ Index_{i,j} = SCI_{i,j}$$

where *FB_Users_i* and *j* represent active Facebook users in countries i and j and *FB_Connections_i,j* represent the amount of friendships between countries i and j. Therefore, the SCI represents the relative probability of a facebook connection between two countries. However, the authors scale the available data so that the Facebook Social Connectedness Index has a minimum value of one and a maximum value of one billion. The higher the social connectedness index value for two countries i and j, the higher their connectedness. If the SCI reaches 1 billion, every facebook user from country I would be connected to every facebook user over country j. (Bailey, Strobel, Wong, Cao, & Kuchler, 2021). The definition in equation (1) denotes that the value of the Social Connectedness Index does not depend on the country's perspective from which it is measured. From the perspective of both country i and country j, the connectedness between them holds equal value.

Furthermore, we use data on regional military expenditures to reveal interstate safety concerns. I choose the dataframe *Military Expenditures* from Stockholm International Peace Research Institute to illustrate the development of regional insecurity measures. The basic assumption behind this time-variant insecurity indicator is as follows: If states feel safe, they do not have to improve their federal spending on the military. On the contrary, if states feel that their neighbourhood’s safety deteriorates, they might increase their military expenditures. That is why this study expresses Military Expenditures in US$ values of 2022 differentiated by continents and respective year to our dataframe (Stockholm International Peace Research Institute, 2024).

Additional information on the relationship between pairs of countries is derived from CEPII. I included dummy variables on the occurrence of colonial relationships, common language and geographical distance (Thierry Mayer & Zignago, 2011). These control variables shall serve as additional indicators for cultural and geographical proximity between two countries.

3.2 Selection & first adjustements

3.2.1 Conflict Data

Gaining insights on the drivers of post-cold war interstate conflicts, I used data from the *Uppsala Conflict Data Program* as baseline for our dependent variable. The authors introduce a range of 21 different steps of conflict escalation, which again can be separated in five linearly escalating hostility levels of conflict. Important to notice: This study slightly adjusts the scale to allow the implementation of logarithmic values in our Shift-Share model. It reaches from 1 (*no militarized conflict*) to 22 (*Join interstate war*) instead of 0-21 (Palmer, et al., 2020). Our model uses *hiact* as dependent variable for interstate conflict intensity. I omit all observations that contain -9 as value for *hiact*.

3.2.2 Social Proximity

To further dig into the relationship between micro relationships and interstate conflict, this study introduces Meta's Social Connectedness Index (SCI) as independent variable representing interstate networking of individuals. The SCI was conducted as snapshot of 2020 user data, gathered in October 2021. However, it is reasonable to assume that the Facebook Social Connectedness Index reflects the overall connectedness of countries, which remains constant over time, as demonstrated by Persson & Tabellini (2009). This also implies that the proximity of countries captured by the Facebook Social Connectedness Index is not related to the availability of the internet. Thus, Facebook data can be used to approximate countries' closeness in years prior to Facebook's founding. Bailey et al. (2021) provide evidence supporting this assumption. Therefore, this paper assumes that the Social Connectedness Index is capable of representing post-cold war-micro relationships sufficiently.

3.2.3 Insecurity Determination

Supplementing the indicator for micro relationships, the SCI, this study further includes an indicator that aims to capture a sense of insecurity perception between states on a continent. For this purpose, we include Military Expenditures (*ME*) per year on a continent level. Intuitively, in years of greater conflicts, states will increase their defence spending to be as prepared as possible if conflicts arise. This indicator shall serve as time variant supplement to the SCI, to add necessary time and location affected tensions. This data is provided by the SIPRI database, which monitors military spending on an annual basis (Stockholm

International Peace Research Institute, 2024). To account for interstate conflicts across continents, this study adopts the indicator to differentiate between the two respective continents, as shown in equation (2)

$$(2)\ MEdiff = 0.5 * ME_{continent_i} + 0.5 * ME_{continent_j}$$

3.2.4 Cultural & geographical proximity

We also include data on colonial relationships, whether the two countries share common languages and the distance between the two capitals. These controls are provided by the CEPII and are necessary to include most basic limitations on interstate conflict, such as geographical distance (Thierry Mayer & Zignago, 2011). It contributes to this study's understanding of proximity between states, which includes cultural and geographical aspects. Since militarized interaction between states requires significant resources in armed forces and logistical capabilities, as well as allies that support transfer of troops and equipment by providing access to their military and civil infrastructure, such as airports or harbours (Martin M. , 2023), distance between capitals might capture a significant amount of conflict intensity variation and supplements the indicators of social proximity.

3.2.5 Historical proximity

Additionally, Cooper (1994) finds that effects of colonial relationships still tend to influence interstate relationships, although the exact impact remains a matter of discussion (Cooper, 1994). My dataset therefore does include a variable on whether there has been some sort of colonial relationship or not in the past, to contribute to recent research on the long-term effect of colonialization and to supplement our understanding of proximity due to historical relationships. Although this does not bring forth sophisticated results on the exact impact of colonial relationship, the indicator can potentially serve as a hint whether the historical component should be the subject of thorough research efforts with regard to its influence on contemporary interstate conflict behaviour.

3.3 Scaling, Adjustments & Scope

Taking a first glance on the distribution of our discrete variables (SCI, ME & distcap) reveals an interesting phenomenon:

Figure 1: Distribution of variables hiact, SCI & distcap via boxplot

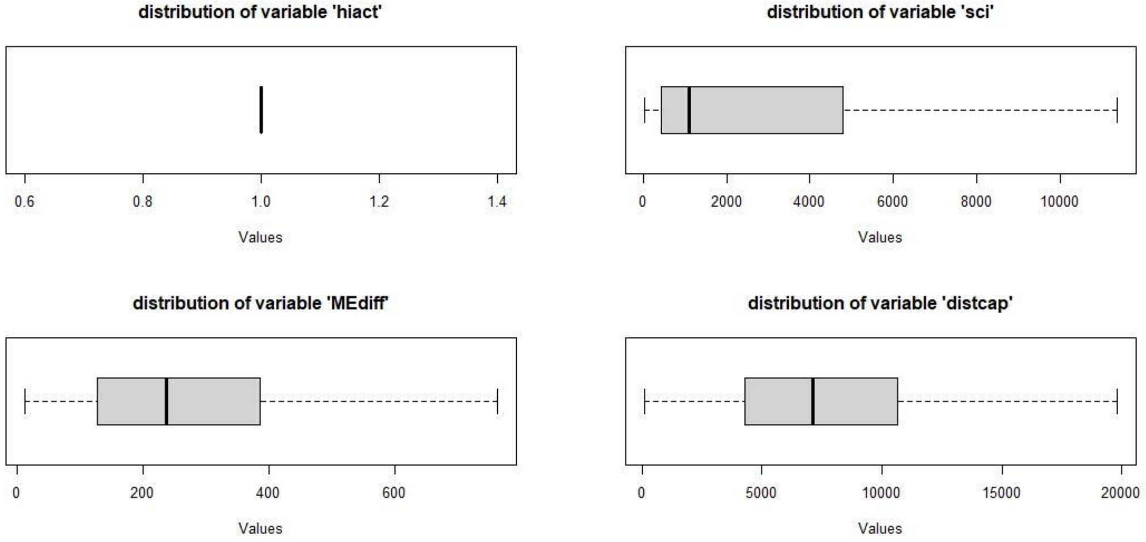

Table 1: Distribution of variables hiact, SCI & distcap via R

	Min.	**Q1**	**Median**	**Mean**	**Q3**	**Max**
hiact	1	1	1	1,016	1	22
SCI	44	428	1,099	18,623.480	4,802	4,249,885
MEdiff	12.519	127.665	237.359	272.745	385.126	985.626
distcap	117.345	4,270.053	7,140.804	7,635.347	10,635.110	19,812.040

The variables are highly right skewed. You can observe this by comparing the calculated mean of the variable with the median in table 1. This is due to the design of our dataframe. We included all countries on a dyadic country-country scheme from 1988-2014

1. … whose micro relationships are monitored by the SCI
2. … where both countries conflict behaviour was monitored in at least one conflict (hiact greater than 1 in at least one observation)

These conditions result in a dataframe that covers not only observations of militarized conflict behaviour, but rather observations on overall international relations between 88 countries in the post-cold war era. The magnitude of our dependent variable *hiact* towards one reveals that these times have mostly been times of peace, as measured by 1 (= non militarized action). The finding in plain English allows to conclude that during that time, the overall majority of countries had been in the state of "*no militarized action*" (= 1) with the absolute majority of countries observed. In fact, over 99% observations of hiact are of value 1. The SCI variable is also strongly right skewed. This is due to some extremely high observations between country pairs. This is evident by comparison of mean and median of the SCI in Table 1. The mean is 18 times higher than the Median. A clear signal that the variable is influenced by strong outliers. E.g., the highest observed value for the SCI is 4,249,885 for Liberia and Sierra Leone. This means, the relative probability that two facebook users from these countries are linked by a facebook relationship is at around 4.2% (Bailey, Strobel, Wong, Cao, & Kuchler, 2021).

Similar characteristics can be observed for the other continuous variables (see Table 1). We attribute this in part to the already presented fact that we choose variables presumably connected to conflict, while designing our dataframe to cover a majority of observations made during non-militarized conflict (peace) times.

To deal with these characteristics, the first approach will be designed to deliver a sense of the impact of our variables. This means, the 2SLS-model will tell, how much variation in conflict levels is explained by our variables. Therefore, this approach includes an Instrumental Variable. The 2SLS design is derived by (Angrist & Pischke, 2008). We scale our continuous variables in a way that shifts the focus from absolute to relative measures. This means, e.g., for the SCI, that we do not measure on a scale of none to absolute connectedness (0-1,000,000,000). Instead, we use relative terms (0-1) with 0 being the lowest observed value of the SCI in the dataframe, 1 being the highest. This adjusts the scale to a relative measure. Dyadic country pairs are now classified, continuing the example of the SCI, on behalf of their connectedness in relation to the highest observed connectedness. This adjustment however, does not change the factor of magnitude towards 0 in our data. To deal with this and allow focussing on the lower bandwidth of our data, we add further weighting to our rescaled relative variables. These measures result in the following variables to be introduced:

$$(3)\ rescaled\ hiact = \left(\frac{hiact - min_{hiact}}{max_{hiact} - min_{hiact}}\right)^{0.1}$$

$$(4)\ rescaled\ sci = \left(\frac{sci - min_{sci}}{max_{sci} - min_{sci}}\right)^{0.1}$$

$$(5)\ rescaled\ MEdiff = \left(\frac{MEdiff - min_{MEdiff}}{max_{MEdiff} - min_{MEdiff}}\right)^{0.1}$$

$$(6)\ rescaled\ distcap = \left(\frac{distcap - min_{distcap}}{max_{distcap} - min_{distcap}}\right)^{0.1}$$

In the rescaling process, including an exponent (0.1) shifts the focus towards lower observations within the dataset. This adjustment aims to emphasize the importance of lower values in our analysis, providing greater attention to the dynamics and characteristics associated with forms of conflict on the lower scale of escalation. This approach allows for a nuanced examination of the data, ensuring that potential insights from lower values are not overlooked. From now on, in regard to variables (3) – (6), we will refer to them as *rsc.hiact*, *rsc.SCI*, *rsc.MEdiff* and *rsc.distcap*.

Analysing the impact of micro-relationships on interstate conflict behaviour presents challenges, particularly due to endogeneity issues. It is possible that aggregated micro-relationships not only influence the escalation of interstate conflict, but that conflict escalation may also impact the aggregated micro-relationships between those states, creating a serious issue of reverse causality. To tackle this challenge, while deepening our understanding of the exact influences of interstate proximity on conflict, we build a Shift-Share-model, following Borusyak, Hull, & Jaravel, (2018) (subsequently called SSA-model or SSA). Doing so, we replace our rescaled variables with logarithmic variables to allow more precise interpretation of our coefficients. The SSA furthermore requires an indicator providing exogenous shocks (shifts) to our independent variable (share). In this study, we have selected an instrument that is endogenous to the dependent variable, namely the highest act of escalation, but exogenous to our independent variable, the SCI. Specifically, we have used aggregated Military Expenditures (in US-$) on continent-level as an indicator for regional sense of insecurity. By using this tool, we aim to enhance the effectiveness of our estimator and reduce the potential bias that may arise from reverse causation. The underlying data sets of the two models are composed as follows.

3.2.1 2SLS data foundation

Table 2: Data foundation of 2SLS model

	Conflict escalation level	**Social proximity layer**	**Regional insecurity indicator**	**Historical proximity layer**	**Geographic proximity layer**	**Cultural proximity layer**
Variable	dependent variable	independent variable	instrument	control	control	control
Name	rsc.hiact	rsc.SCI	rsc.MEdiff	colony	rsc_distcap	comlang_ ethno
Characteristic	cathegoric	continous	continous	dummy	continous	dummy
Range	[0:1]	[0:1]	[0:1]	[0,1]	[0:1]	[0,1]

3.2.2 Shift-Share data foundation

Table 3: Data foundation of Shift-Share model

	Conflict escalation level	**Shift-Share Instrument**		**Historical proximity layer**	**Geographic proximity layer**	**Cultural proximity layer**
Variable	dependent variable	Share-variable	Shift-variable	control	control	control
Name	Log(hiact)	Log(SCI)	Log(ME_diff)	colony	Log(distcap)	comlang_ ethno
Characteristic	cathegoric	continous	continous	dummy	continous	dummy
Range	[0:1]	[3,784: 15,262]	[4.765: 9.894045]	[0,1]	[0.927:1,93]	[0,1]

3.3 OLS

A simple, yet popular way to estimate the effect of micro relationships on interstate conflict would be to use an *Ordinary Least Squares* design as derived multiple times (Angrist & Pischke, 2008; Wooldridge, 2019) and first published by Carl Friedrich Gauss (1809) & Adrien-Marie Legendre (1805). An OLS design minimizes the sum of the squared differences between observed values and those predicted by a linear model. This minimization provides the best fit line for the given data. OLS is particularly effective for linear regression, a technique widely used across various fields such as economics, finance, and the natural sciences, to predict outcomes and determine significant influencing factors (Wooldridge, 2019). In our case, an OLS design could look like this:

$$(7)\ hiact_{i,j,t} = \alpha + \beta_1 * SCI_{i,j} + \beta_2 * MEdiff_{i,j,t} + \beta_3 * colony_{i,j} + \beta_4 * distcap_{i,j} + \beta_5 * comlang_{i,j} + \varepsilon_{i,j,t}$$

Where *hiact* serves as dependent variable, the impact of our independent variables and controls is captured by the β and other influences are captured in the error term $\varepsilon_{i,j,t}$.

This design however, does not suit our data due to two possible endogeneity issues (Angrist & Pischke, 2008). OLS estimations are vulnerable to reverse causality issues. Reverse causality occurs when the relationship between dependent variables and independent variables are two-fold. In this case, independent variables do not only influence dependent variables, but also vice versa. This triggers serious inaccuracies in the estimation (Wooldridge, 2019). In our case, we can not reject the hypothesis that micro relationships might influence interstate conflict escalation, but that interstate conflict escalation also has an impact on micro relationships. If, e.g., two states engage in advanced militarized conflict with each other, aggregated micro relationships between participating states might suffer due to travel restrictions, information campaigns of each state against the other, trust erosion, communication barriers, et cetera. Phenomena, that do occur in militarized conflicts (Martin M. , 2023).

Additionally, our OLS-research framework, as derived in equation (7) might be vulnerable to omitted vairables. Omitted variables, or multicollinearity issues, occur if the variation within a dependent variable is related to the variation in another dependent variable (Angrist & Pischke, 2008), as illustrated in equation (8):

$$(8)\ Cov[x_1, x_2] \neq 0$$

Advanced multicollinearity results in biased estimates. In our case, the indicator for cultural proximity, presence of a common language (*comlang*) will most likely impact the indicator for cultural procimity, the SCI, since a common language between states might influence the occurrence of micro relationships across these states. The same accounts for colonial relationships, which most likely have an impact on the SCI due to migration and displacement of people from colonized countries to colonizers and counter streams (Engerman, 1986).

These factors prevent compliance with the Gauss-Markov theorem (Gauss, 1823) and therefore completely hinder the efficacy of OLS designs in our research environment.

3.4 2SLS

One way to address issues like these, is to use instrumental variables. Our 2SLS model is largely based on derivations by Wooldridge (2019) and Borusyak et al. (2018).

2SLS estimations are widely recognized as proper identification methodology to estimate linear equations in applied economics where reverse causality or endogeneity issues arise (Wooldridge, 2019). 2SLS models depend on instruments that provide additional information on the independent variable in question. The idea is to choose an instrument *z* that is related to the dependent variable *y* but unrelated to the independent variable so that

$$(9) \qquad Cov[z, y] \neq 0$$

$$(10) \quad Cov[z, x_1] = 0$$

holds (Wooldridge, 2019). This introduces exogenous variation to our dependent variable, solves the reverse causality issues and implements further variation as we choose a time variant instrument. Since we observe data on interstate conflict escalation, a suitable instrument to our model would be an indicator that does influence conflict, but on a similar level for a large amount of individuals, hence states. An indicator that captures a sense of insecurity that accounts for many states but stays exogenous to micro relationships between states. This study chooses aggregated Military Expenditures on continental scale. The intuition behind that is to provide an indicator that captures a sense of insecurity among a larger group of observed individuals which is persistent to individual outlayers. On the latter point: If one country improves their military expenditures, the effect on the aggregated expenditures on continental scale will hardly be affected. The continental scale further provides exogenous variation that might not be correlated to bilateral proximity indicators as implemented in this study so that equation (10) holds. These requirements result in the following model:

$$(11) \quad rsc.hiact_{i,j,t} =$$
$$\alpha_1 + \beta_1 * X_{i,j,t} + \beta_2 * colony_{i,j} + \beta_3 * rsc.distcap_{i,j} + \beta_4 * comlang_{i,j} + \varepsilon_{i,j,t}$$

with

$$(12) \quad X_{i,j,t} = \alpha_2 + \delta_1 * rsc.SCI_{i,j} + \delta_2 * rsc.MEdiff_{i,j,t} + \mu_{i,j,t}$$

Our model uses the highest act of escalation (*hiact*) as dependent variable, and regional military expenditures as insecurity indicator (*MEdiff*) as instruments, dummy variables whether the two parties share a common language or have a colonial past (*comlang_ethno* & *colony*) and a geographic correctiv (*distcap*) as controls. All rescaled as derived in equations (3)-(6). Where

- $rsc.hiact_{i,j,t}$ is the highest level of conflict escalation between countries i and j in year t and serves as dependent variabel
- $X_{i,j,t}$ is a variable derived as the predicted value from a regression where $rsc.SCI_{i,j}$ is the dependent variable, and $rsc.MEdiff_{i,j,t}$ serves as the main instrument. This predicted value of $rsc.SCI_{i,j}$ is then used in the second stage regression to assess its impact on the dependent variable, $rsc.hiact_{i,j,t}$ and error term $\mu_{i,j,t}$

- $colony_{i,j}$ captures historical proximity per dummy variable on the colonial history
- $comlang_{i,j}$ captures cultural proximity per dummy variable whether both states share a common language
- $rsc.distcap_{i,j}$ captures geographical proximity as per distance between the capitals of both sides
- $\epsilon_{i,j,t}$ is the residual, accounting for random variability in conflict escalation not explained by the model.
- α_1 and α_2 are constants

3.5 SSA

Our second model uses a shift-share design, as established by Timothy Bartik (1991) to combine the time invariant SCI and the time variant ME to a shift-share instrument that shall capture the influence of micro relationships on the dependent variable *hiact*, supplemented by controls on common language (*com_lang*), distance between capitals (*distcap*) and dummy on colonial relationships (*colony*) on a dyadic country-country scheme between 1989 and 2014, only including interstate conflicts between two states.

This approach is based on the work of Borusyak, Hull, & Jaravel, (2018), who provide a comprehensive framework for employing shift-share designs in economic research. Their paper elucidates the robustness of shift-share instruments in isolating regional impacts from national shocks, making it a seminal work for any research employing this analytical design. The authors differentiate between two basic methods: the basic Shift-Share Instrument design and the quasi-experimental shift-share design. The former is a standard framework for using shift-share instruments, which combine observed shocks with unit-specific weights to estimate regional effects of national trends or policies. The latter, a more advanced approach, introduces methodological enhancements that ensure the exogeneity of these instruments through rigorous testing and validation processes (Borusyak, Hull, & Jaravel, 2018). This study chooses a basic shift share design, since the model does not depend on exogenous weighting scales, which are often used in use-cases of labour-industrial analysis (e.g. Elsby & Şahin, 2013). The design of Borusyak et al. (2018) is well suited for this case and adopted as following. I choose an instrument

$$(13) \qquad z_{i,j,t} = \sum_{i|j} s_{i,j} g_{i,j,t}$$

that captures regional shocks g at time t that affect both participants i & j, as well as the share of the dependent variable of both participants s. The adoption proceeds as following:

$$(14) \quad z_{i,j,t} = \sum_{i|j} \log(sci_{i,j}) * \log(MEdiff_{i,j,t})$$

I now seek to estimate the causal effects β_1 of the linear model

$$(15) \quad y_{i,j,t} = \alpha + \beta_1 * z_{i,j,t} + \beta_n * w_n + \epsilon_{i,j,t}$$

Where

- $y_{i,j,t}$ is our outcome
- α is our constant
- β_1 is the estimate for our shift-share instrument $z_{i,j,t}$
- β_n are our estimates for a set of w_n control variables and
- $\epsilon_{i,j,t}$ is the residual that captures all other not-observed influences on the outcome and is defined to be orthogonal to our set of control vectors w_n

More about the properties of the orthogonality condition later in this chapter. This model implements logarithmic values to our continuous variables to later allow more precise interpretation of my results. My model is specified as follows:

$$(16) \quad \mathrm{Log}(hiact_{i,j,t}) =$$

$$\alpha + \beta_1 * z_{i,j,t} + \beta_2 * colony_{i,j} + \beta_3 * log\ (distcap_{i,j}) + \beta_4 * comlang_{i,j} + \epsilon_{i,j,t}$$

Where

- $log(hiact_{i,j,t})$ is the highest level of conflict escalation between countries i and j in year t in logarithmic terms
- $z_{i,j,t}$ is the shift-share instrument combining logarithmic social connectedness and logarithmic continental military expenditure
- $colony_{i,j}$ includes the dummy variable on the colonial history
- $comlang_{i,j}$ includes the dummy variable whether both states share a common language
- $log\ (distcap_{i,j})$ captures the geographical distance between the capitals of both sides
- $\epsilon_{i,j,t}$ is the residual, accounting for random variability in conflict escalation not explained by the model.

Shift-share models, as derived on the baseline of Borusyak et al. (2018) do not rely on the popular assumption of independent and identically-distributed (iid) data. Instead, the shift-share designs usually rely on the full-data moment condition to capture the orthogonality of the shift-share instrument (Borusyak, Hull, & Jaravel, 2018). Since this study does not include weights in the shift-share instruments, the full-data moment condition equals a conventional condition of iid. This means that

$$(17) \quad E\left[\sum_t z_{i,j,t}\epsilon_{i,j,t}\right] = 0$$

must hold (Borusyak, Hull, & Jaravel, 2018). In this case, the conventional condition of iid is

$$(18) \quad E\left[\sum_t z_{i,j,t}\epsilon_{i,j,t}\right] = -6.384864e - 19$$

To check whether the deviation of zero is significant enough to reject the null hypothesis, we conduct a t-test which results in an effective p-value of 1. This allows us to reject concerns over endogeneity issues of our instruments. The result is attached in the Appendix.

This approach allows us to account for interdependencies and shared variance within dyads, a common occurrence in data involving interactions between specific pairs, such as between countries in international relations studies.

4 Results

4.1 2SLS

	hiact	
	(1)	**(2)**
X	0.002***	-0.018***
	(0.002)	(0.002)
colony		-0.0004
		(0.001)
rsc.distcap		-0.04***
		(0.001)
comlang_ethno		0.002***
		(0.0002)

Constant	-0.00002	0.045***
	(0.001)	(0.001)
Observations	209,942	209,942
R2	0.00001	0.007
Adj. R2	0.00000	0.007
Residual Std. Error	0.024	0.024
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 4: 2SLS model with rescaled values as derived in equations (3)-(6). Source: See Chapter 'Data'.

Our IV model does find some significant influences on two-party interstate conflict behaviour from 1989 to 2014, as table 4 shows. Our does have a significant impact on conflict, lowering escalation levels by -0.018 units(Table 3, column 2, row 1). The coefficient of -0.018 for the instrument suggests that a one-unit increase in the instrument is associated with a decrease of 0.018 units in *rsc.hiact*, all else being equal. The colonial relationship does not seem to have any impact on conflicts. However, the distance between states does have a negative influence on conflict escalation, as our estimator for the variable *rsc.distcap* shows. The coefficient of -0.04 for *rsc.distcap* suggests that a one-unit increase in the *rsc.distcap* is associated with a decrease of 0.04 units in *rsc.hiact*, all else being equal. This finding is quite intuitive, since larger distances complicate militarized conflict, since the transport of military forces across larger distances requires high ressources, supporting allies and induces higher costs to the engagement (Martin M. , 2023). Finally, our model suggests a significant positive relationship between conflict escalation and cultural proximity as derived by this study. The coefficient of 0.002 suggests that the appearance of a common language enhances conflict escalation between the associated states by 0.002.

This study implements a standard deviation approach to explain how much variation of our dependent varible *rsc.hiact* depends on our variables. This measurement is conducted as following:

$$var_{rsc.hiact} = \frac{(\beta_k * \delta_k)}{\delta_{rsc.hiact}} \quad (19)$$

With

- $var_{rsc.hiact\,|\,k}$ = being the amount of variation of rsc.hiact that is explained by variable k
- β_k = coefficient of variable k
- δ_k = standard deviation of variable k
- $\delta_{rsc.hiact}$ = standard deviation of *rsc.hiact*

Variable	**Variation of rsc.hiact explained**
$X_{i,j,t}$	$\frac{\left(\beta_{X_{i,j,t}} * \delta_{X_{i,j,t}}\right)}{\delta_{rsc.hiact}} = -\frac{0.018 * 0.03665112}{0.03270099} \approx -0.0201$
$colony_{i,j}$	$-\frac{0.0004 * 0.1138208}{0.03270099} \approx -0.0014$
$rsc.ditscap_{i,j}$	$-\frac{-0.04 * 0.06923193}{0.03270099} \approx 0.0847$
$comlang\_ethno_{i,j}$	$\frac{0.002 * 0.3519163}{0.03270099} \approx 0.0215$

Table 5: Variation explained by our 2SLS model per variable.

Measuring the amount of variation in conflict levels across our panel as explained with equation (19) shows that our variable X that includes the SCI and the regional insecurity indicator as instruments explains roughly 2% of variation in conflict levels (see Table 5). Our indicator for cultural proximity explains 2% of variation as well. Furthermore, our geographical proximity indicator distcap explains almost four times as much variation (8%) in conflict levels. A clear indicator that geographical proximity does have strong influence on interstate conflict behaviour.

The results suggest a complex relationship between micro relationships and corresponding interstate conflict-behaviour. To further deepen the understanding of the drivers of interstate conflicts, this study implements a SSA model with logarithmic values.

4.2 SSA

	hiact	
	(1)	**(2)**
Log(SCI)	0.002*** (0.0001)	0.001*** (0.0001)
colony		-0.002 (0.001)

Log(distcap)		-0.009*** (0.0003)
comlang_ethno		0.004*** (0.001)
Constant	-0.021*** (0.001)	0.070*** (0.003)
Observations	209,942	209,942
R2	0.004	0.008
Adj. R2	0.004	0.008
Residual Std. Error	0.033 (df = 209940)	0.033 (df = 209937)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 6: Sift-Share model with rescaled values as derived in euqations (3)-(6). Source: See Chapter 'Data'.

This study earlier unveiled a certain trend of our original parameters in an 2SLS model, as Table 3 shows. Using logarithmic variables in our SSA model sharpens our understanding of the actual dynamics once again. The negative influence of the SCI on conflict escalation levels does not proceed. Instead, a one unit increase in the SCI triggers a 0.1 % increase in conflict escalation across the sample. This result is highly significant at the 0.01 level. Our social proximity indicator therefore increases conflict levels on a significant level, as does our cultural proximity indicator. If two countries share a common language, conflict escalation increases by 0.4% on average. Furthermore, as stated by many authors including Martin (2023), my model supports the negative relationship of distance between countries and conflict escalations. 1% increase in the distance between capitals reduces conflict escalation by 0.9%. Geographical proximity is therefore the only one of our proximity terms that reduces conflict levels. This result is also significant at the 0.01 level. The historical proximity indicator is the only variable that has no significant influence on our conflict variable.

Overall, while the 2SLS model reveals certain trends in my data, the SSA model refines the picture. While the identified trends for historical, geographical, and cultural proximity have been confirmed in significance and direction, the result for social proximity has changed. Implementing the logarithmic-based SSA model has therefore been crucial to identify causal relationships of proximity layers on conflict escalation mechanisms on bilateral interstate level.

5 Discussion

This study contributed to the field of interstate conflict research and international relations. The presented models are based on a dataframe covering bilateral relations between 88 countries during 27 years. However, the included variables of this study only allow interpretation under certain conditions, of which some have been acknowledged before, but this chapter shall serve as guidance on interpretation and challenges with the chosen research framework.

First of all, as shown in Chapter 4, the formidable right skewness of our discrete variables (see Table 1 & 2) has complicated our measures. Our dataset includes annual data on international relations between 88 states over 27 years, resulting in almost a quarter million observations. Right skewness of our data suggests during this time, that the overwhelming amount of annual data on bilateral relations of the included states have been gathered during times of no military escalation, hence, peace. Intuitively, to gain knowledge on conflict dynamics, this study focuses on the variation in bilateral relations that happen above that level of no militarized conflict (or peace) and therefore introduced rescaled variables, as derived in equations (3) – (6). Unfortunately, it prohibits absolute interpretations of the derived estimates, as the results are derived by rescaled variables, rather than absolute. We can still acknowledge trends, whether variables increase or decrease our dependent variable. But we can not proceed by stating absolute values by which our indicators increase or decrease the absolute conflict escalation scale in our 2SLS model. Upcoming research could address this limitation by only observing times of conflict, rather than absolute timeframes, dominated by times of no escalation. The trend difference for the social proximity indicator may also be partly caused by the different focus of the 2SLS model and the SSA model. Where the 2SLS model contributes to our right skewed data and therefore includes rescaling, the SSA model does not shift the focus on our variables to allow more precise findings. This means, the shift of the SCI from reducing conflict escalation levels to benefit such might result partly from a focus that weighs observations differently.

Some rather prominent limitations lie in the design of the dataframe. The dyadic scheme is well suited to bilateral observations. Conflicts, where multiple parties engage in militarized actions are not included in this scheme. All findings on the influences of interstate conflict are limited to two-state occurrences. Another limitation related to this is the dimension of conflict included in the analysis, which only includes conflicts between two states that fulfil certain conditions as derived before (*see Chapter 3.1*). But large-scale militarized conflicts do not only happen between two (or multiple) states, but often between states and non-states, or state-like actors. Sometimes in the form of civil conflicts, where for example, governments fight rebellions, sometimes in the form of state-state-like conflicts, where for example governments engage in a conflict with states that are not internationally recognized.

Relating to the context of micro-relationships: META's Social Connectedness Index is well recognized in international trade research (Bailey, et al., 2021) and has already served as determinant for social capital in political science reviews (Chetty, et al., 2022). However, it does come with certain limitations regarding the introduction of noise for reasons of privacy of affected users (Bailey, Strobel, Wong, Cao, & Kuchler, 2021), which makes the indicator less precise and more severe: the selection of countries. Afghanistan, Western Sahara, China, Cuba, Iraq, Israel, Iran, North Korea, Russia, Syria, Somalia, South Sudan, Sudan, Venezuela, Yemen and other state-like regions like Gaza are not included due to restrictions to META's commercial activities in those countries (Bailey, Strobel, Wong, Cao, & Kuchler, 2021). Without conducting any analysis whether these countries do or do not engage frequently in global interstate conflict, only their sheer number alone limits the analytical power of this study. However, the authors emphasize that they only included data from countries, where META's corporate activity, and therefore the regulatory framework of the respective state, does allow data extraction under academic standards (Bailey, Strobel, Wong, Cao, & Kuchler, 2021). Under the assumption that those standards are violated more often by autocratic systems rather than democratic systems, the exclusion might just remove significant actors in global conflicts, as several papers suggest a positive relationship between international armed conflicts and autocratic states (e.g. (Conrad, 1905)). If the assumption over the autocratic tendencies in excluded countries holds, other indicators for interstate micro relationships can be considered to supplement research on the topic, such as flight connections or internet traffic.

Furthermore, while the implementation of dummy variables on the existence of common languages between conflict parties or whether conflict parties have had a colonial relationship should lead to further research, the indicators remain simple and do not further differentiate between different forms of colonial relationships nor do they differentiate between still proceeding relationships or dissolved ones.

On behalf of the empirical strategy, a central assumption is the exogeneity between the SCI and the regional insecurity indicator, military expenditures on continental level. While at the point of writing, literature on possible connections between the two variables pends, future research efforts could prove otherwise.

This study focused on the influence of proximity on interstate conflict. However, the approach also allows to implement other potential drivers on macro-level conflict. Michael Doyle (1993) finds that democracies do not tend to fight each other. This is due to shared norms, common mechanisms for conflict resolution & institutions (Doyle, 1983). This finding, often referred to as “democratic peace theory” has been much discussed in international relations research and provides an interesting impulse to our research framework: the implementation of a political proximity indicator, including variables determining the political system to our models. Under the assumption that micro relationships present a significant influence on voters decision-making process and considering my results, the level of democratisation of the political system of a state should increase interstate conflict behaviour. One recognized data source on the democratic level is the Polity5 project, published annually by the Center of Systemic Peace. This project assigns a Polity Score to certain national states that captures the democratic level of the political system on a scale between -10 (full autocracy) to 10 (full democracy) (Center for Systemic Peace, 2024). A potential indicator to include in further research efforts on the subject of interstate conflict.

6 Conclusion

This research ventured into an examination of the dynamics underlying interstate conflicts, adapting the baseline research framework from Allport's contact theory (1954), which traditionally analyzes influences on intergroup relationships, to international relations. Through this lens, the study systematically investigated how different layers of proximity between nations influence the escalation of bilateral conflicts. Using comprehensive datasets from institutions like the Stockholm International Peace Research Institute, Uppsala University, and META, the analysis employed robust 2SLS & Shift-Share models to delve into this intricate subject matter.

The findings of this thesis reveal several important findings on the influence of different proximity layers on conflict between states. Our SSA-model finds that a 1 % increase in social proximity, as represented by the SCI, does increase conflict levels by 0.1%. Cultural proximity, as represented by the occurrence of a common language increases conflict levels across our panel by 0.4%. Geographical distance reduces conflict levels by 0.9%. In return, if geographical distance reduces conflict escalation significantly, geographical proximity increases conflict probability. The results therefore added to the findings of Caselli, Morelli, & Rohner (2015), which included a weighting of economic incentives to their geographical estimator and found it to be a significant driver of bilateral conflict probability. Our estimation on historical proximity did not reveal any significant influences on conflict. This means, three of four indicators of interstate proximity increased conflict escalation significantly. This may be interpreted as a hint that Allport contact theory (1954) does not hold on to a macro level as analysed in our framework of four layers of proximity.

Additionally, this study's 2SLS-model reveals that the social proximity layer, as well as our cultural proximity layer, both explain roughly 2% of conflict variation across our panel (see Table 5). The geographical proximity layer explains four times as much variation in conflict levels, more than 8%.

Our findings are in line with numerous other research studies on the impact of different layers of interstate proximity on militarized conflict: Therefore, we must reject Eisenhower's thesis quote at the beginning. As great as its contribution to the resolution of military conflicts and the creation of sustainable peace in Europe may have been, this study rejects his assessment on the calming impact of micro relationships on global interstate conflict.

7 Appendix

7.1 Full-moment data condition of SS instrument

Testing the orthogonality of our Shift-Share instrument $z_{i,j,t}$ to our error term $\epsilon_{i,j,t}$ requires a full-moment data condition to hold. This means we conduct Hypothesis testing as conducted in Angrist & Pischke (2008). With

$$(20) \quad H_0 = E\left[\sum_t z_{i,j,t}\epsilon_{i,j,t}\right] = 0$$

$$(21) \quad H_A = E\left[\sum_t z_{i,j,t}\epsilon_{i,j,t}\right] \neq 0$$

7.1.1 Results

T-Statistics	-1.6317e-11
Degrees of Freedom	209941
P-Value	1.0
95%-Confidence Interval	[-0.0000734, 0.0000734]
Mean of the product	-6.113822e-16

7.1.2 Interpretation

The t-test results indicate that the mean of the product between the instrumental variable and the residuals is statistically indistinguishable from zero. The very high p-value (1.0) and the confidence interval containing zero strongly support the null hypothesis. This outcome confirms the orthogonality of the instrumental variable to the error terms, validating the instrument's exogeneity and appropriateness for use in the 2SLS regression model.

7.2 F-Statistic in 2SLS model

The first-stage F-statistic is a critical measure in Instrumental Variables (2SLS) regression, used to assess the strength of the instruments employed. It tests the null hypothesis that the instruments are jointly insignificant in explaining the variation in the endogenous explanatory variables. A widely accepted rule of thumb is that an F-statistic value greater than 10 indicates that the instruments are sufficiently strong to provide reliable 2SLS estimates, mitigating the concerns of weak instrument bias that can lead to inconsistent estimators (Angrist & Pischke, 2008).

$$(22) \quad H_0 = F = \frac{\left(\frac{R^2}{k}\right)}{\frac{1-R^2}{n-k-1}} < 10$$

$$(23) \quad H_A = F = \frac{\left(\frac{R^2}{k}\right)}{\frac{1-R^2}{n-k-1}} > 10$$

7.2.1 Results

F-Statistic	13,651.5
Degrees of Freedom	209937
P-value	<0.0001

7.2.2 Interpretation

The F-statistic obtained from the 2SLS Model 13,651.5, which far exceeds the threshold of 10, suggesting that the instruments employed are very strong. The extremely low p-value (< 0.0001) further rejects the null hypothesis of the instruments being weak. This implies that the instruments used are valid and effectively explain the variation in the endogenous variable, leading to reliable and consistent 2SLS estimates.

By demonstrating a robust instrument strength via the F-statistic, the concerns regarding the potential bias and inconsistency often associated with weak instruments in 2SLS estimation are mitigated. Therefore, the results substantiate the reliability of the 2SLS approach adopted in this analysis.

7.3. Dependent Variable composition

Variable hiact from Dataset MIDB 5.0

Highest act of escalation [hiact]

- 1 = No militarized action [1]
- 2 = Threat to use force [2]
- 3 = Threat to blockade [2]
- 4 = Threat to occupy territory [2]
- 5 = Threat to declare war [2]
- 6 = Threat to use CBR weapons [2]
- 7 = Threat to join war [2]
- 8 = Show of force [3]
- 9 = Alert [3]
- 10 = Nuclear alert [3]
- 11 = Mobilization [3]
- 12 = Fortify border [3]
- 13 = Border violation [3]
- 14 = Blockade [4]
- 15 = Occupation of territory [4]
- 16 = Seizure [4]
- 17 = Attack [4]
- 18 = Clash [4]
- 19 = Declaration of war [4]
- 20 = Use of CBR weapons [4]
- 21 = Begin interstate war [5]
- 22 = Join interstate war [5]
- -9 = Missing [-9]

Hostility level [hostlev]

- [1] = No militarized action

- [2] = Threat to use force
- [3] = Display of force
- [4] = Use of force
- [5] = War …

(Palmer, et al., 2020)

8 Bibliography

Allport, G. W. (1954). *The Nature of Prejudice.* New York: Addison-Wesley.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics | An Empiricist's compagnion.* Economics Books, Princeton University Press.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R., & Stroebel, J. (2021). International trade and social connectedness. *Journal of International Economics*.

Bailey, Strobel, J., Wong, A., Cao, R., & Kuchler, T. (2021, 03 04). Social Connectedness: Measurements, Determinants, and Effects. Retrieved from https://dataforgood.facebook.com/dfg/docs/methodology-social-connectedness-index

Bartik, T. J. (1991). Who Benefits from State and Local Economic Development Policies? *Kalamazoo: W.E. Upjohn Institute*.

Borusyak, K., Hull, P., & Jaravel, X. (2018). *Quasi-Experimental Shiftshare Research Designs.* Cambridge: NATIONAL BUREAU OF ECONOMIC RESEARCH.

Bosquet, C., Combes, P.-P., Henry, E., & Mayer, T. (2022). Peer Effects in Academic Research: Senders and Receivers. *The Economic Journal*, 2644-2673.

Caselli, F., Morelli, M., & Rohner, D. (2015). The Geography of Interstate Resource Wars. *Quarterly Journal of Economics*, 267-315.

Center for Systemic Peace. (2024). Polity5: Regime Authority Characteristics and Transitions Datasets.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., . . . Wernerfelt, N. (2022). Social capital I: measurement and associations with economic mobility. *Nature*, 108-121.

Conrad, J. (1905). Autocracy and War. *The North American Review*, 33-55.

Cooper, F. (1994). Conflict and Connection: Rethinking Colonial African History. *The American Historical Review*, 1516-1545.

Department of Peace and Conflict Research . (2023, 03 13). *UCDP Background*. Retrieved from https://www.pcr.uu.se/research/ucdp/about-ucdp/ucdp-background/

Doyle, M. W. (1983). Kant, Liberal Legacies, and Foreign Affairs. *Philosophy & Public Affairs*, 323-353.

Eisenhower, D. D. (1954). *address to a meeting of the National Council of Catholic Women.* New York Times.

Elsby, M. W., & Şahin, A. (2013). *The Decline of the U.S. Labor Share.* Brookings Papers on Economic Activity.

Engerman, S. L. (1986). *Servants to slaves to servants: contract labour and European expansion.* Martinus Nijhoff Publishers.

Esteban-Marquillas, J. (1972). I. A reinterpretation of shift-share analysis. *Regional and Urban Economics*, 249-255.

Fox, J., Kleiber, C., Zeileis, A., Kuschnig, N., & Tam, R. C. (2024). *Instrumental-Variables Regression by '2SLS', '2SM', or '2SMM',.* CRAN.

Gauss, C. F. (1823). *Theoria combinationis observationum erroribus minimis obnoxiae.* Göttingen.

Hess, G. D., & Orphanides, A. (2001). War and Democracy. *Journal of Political Economy*, 776-811.

Martin, M. (2023). *How to fight a war.* C Hurst & Co Publishers Ltd .

Martin, P., Mayer, T., & Thoenig, M. (2008). Make Trade Not War? *The Review of Economics Studies*, 965-900.

Palmer, G., McManus, R. W., D'Orazio, V., Kenwick, M. R., Karstens, M., Bloch, C., . . . Soules, M. J. (2020). *Codebook for the Militarized Interstate Dispute Data, Version 5.0.* Uppsala: Correlates of War Project.

Persson, T., & Tabellini, G. (2009). Democratic Capital: The Nexus of Political and Economic Change. *American Economic Journal: Macroeconomics*, 88-126.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annu. Rev. Psychology*, 65-85.

Sims, V. M., & Patrick, J. R. (1934). Attitude toward the Negro of Northern and Southern College Students. *The Journal of Social Psychology*, 192-204.

Skaperdas, S., & Syropoulos, C. (2001). Guns, Butter, and Openness: On the Relationship Between Security and Trade. *Amercian Economic Review*, 353-358.

Stockholm International Peace Research Institute . (2024, 03 28). *SIPRI Military Expenditure Database.* Retrieved from https://www.sipri.org/databases/milex

THE CONVENANT OF THE LEAGUE OF NATIONS. (1919). *Report of the comission on the League of Nations.*

Thierry Mayer, & Zignago, S. (2011). Notes on CEPII's distances measures: The GeoDist database . *CEPII Working Paper*.

Wooldridge, J. (2019). *Introductory Econometrics.* Cengage Learning.