Research Article – Meta-Analysis

SAGE Open
October-December 2024: 1–19
© The Author(s) 2024
DOI: 10.1177/21582440241299976
journals.sagepub.com/home/sgo
Sage

A Meta-Analysis of the Impact of External Interventions on Conflict Intensity: Navigating Heterogeneity and Unveiling Genuine Effects

Binyam Afewerk Demena[1], Amar Anwar[2], and Syed Mansoob Murshed[1,3]

Abstract

We investigate 833 reported estimates detailing the impact of external interventions aimed at alleviating civil war on conflict intensity. This investigation encompasses 34 studies conducted between 1996 and 2020, primarily focusing on African countries. While the average reported effect is both negative and statistically significant, our analysis reveals substantial divergence in the results. We apply meta-regression analysis to examine the sources of this heterogeneity. Our main findings are as follows. First, differences in data, intervention targets, conflict intensity measures, and the publication year of the primary studies account for the observed heterogeneity in reported estimates. Second, we find no evidence of publication selection bias that aligns with prior beliefs, theoretical expectations, or statistical significance. Third, after considering potential sources of heterogeneity and publication bias, the overall genuine effect of external intervention remains negative and statistically significant. This implies that external intervention efforts do mitigate conflict intensity, albeit with a small magnitude of the statistical significance coefficient. For policy purposes, external interventions are likely to have a modest effect.

Plain language summary

Do external interventions restrain conflict intensity? A meta-analysis of evidence on heterogeneity and genuine effects

We looked at whether external interventions affect the intensity of conflicts predominantly in African countries for studies published from 1996 to 2020. It seems like external interventions usually reduce conflict intensity, but the results from different studies vary a lot. We studied 833 reported results from 34 studies to understand why. Overall, the average effect of these interventions is negative, meaning they do reduce conflict, but the results differ because of things like different data, targets of the interventions, and how conflict intensity is measured. We also found that there's no publication selection bias in how studies are published. After considering all these factors, we still see a statistically significant negative effect, meaning external interventions do help reduce conflict intensity, but the impact is modest.

Keywords

external intervention, conflict intensity, genuine underlying effect, heterogeneity, meta-analysis

[1]International Institute of Social Studies, Erasmus University Rotterdam, The Hague, The Netherlands
[2]Shannon School of Business, Cape Breton University, Sydney, NS, Canada
[3]Coventry University, UK

Amar Anwar is now affiliated to International SEPT Competence Center, Universität Leipzig, Germany.

Corresponding Author:
Binyam Afewerk Demena, International Institute of Social Studies, Erasmus University Rotterdam, Kortenaerkade 12, 2518 AX, The Hague, The Netherlands.
Email: demena@iss.nl

Data Availability Statement included at the end of the article

Introduction

Conflict and civil war are a major cause of development failure in low- and middle-income countries. Consequently, prioritizing conflict de-escalation and mitigation becomes imperative in the development efforts, aiming to foster progress and prosperity in the world's developing countries. One suggested mechanism to mitigate civil war, especially its intensity, is external intervention by foreign powers. These intervenors may include a neighbouring country, great powers, or multilateral organizations such as the United Nations (UN). From the standpoint of external powers, mitigating the intensity of civil war also serves a non-altruistic objective, because lowering civil war intensity stems the tide of refugees and obviates the need for deeper and more costly humanitarian interventions in the future when conflict intensifies. The 20-year external intervention in the Afghan civil war serves as an illustrative example, beginning with substantial resources but eventually petered out into cheap talk (refer to Appendix A for the theoretical discussion). The ongoing Russian-Ukraine war and the recent internal conflict in Sudan further underscore the topical nature of external interventions. In these instances, the role of external powers in influencing conflict dynamics becomes even more pronounced and relevant.

We conduct a comprehensive meta-analysis of the existing empirical literature on external interventions in civil wars and their effect on conflict intensity, chiefly measured through battle deaths. We have documented the large numbers of empirical studies conducted until 2020 predominantly for African countries. This thorough review has identified and analysed 833 reported external intervention estimates on conflict intensity from 34 studies. Figure 1 visualizes the extent of disagreement both in terms of the direction of the effect size and their statistical significance. Slightly more than one-third of the estimates (35%) reported a negative and statistically significant impact on conflict intensity. In contrast, about one-fifth of the reported estimates find a positive and statistically significant impact. In the remaining 44%, the evidence is mixed, but statistically insignificant. Despite the voluminous literature exploring the associations between external intervention and conflict intensity, along with international relations policy efforts aimed at promoting peaceful conflict resolution, the empirical studies yield substantially divergent results. The first motivation of this study is to scrutinize whether external intervention actually reduces conflict intensity.

Figure 2 portrays the development of the reported estimates overtime in individual empirical studies, organized by the year of publication in Google Scholar. Dixon's seminal paper in 1996 initially indicated a negative impact of external interventions on conflict intensity.

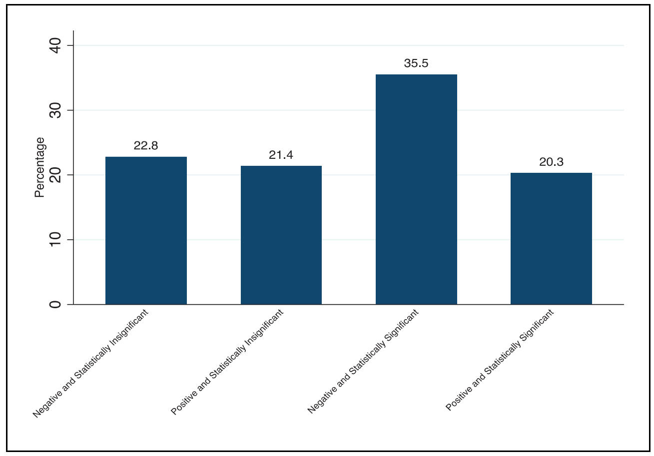

Figure 1. Sign and significance of external intervention impact on conflict intensity ($N = 833$).

Subsequent publication in 2004 introduced evidence of both negative and positive impacts. Despite a wealth of studies in the last decade, substantial variations persist in the estimated parameters. A second motivation of this study is, therefore, to explore the reasons behind this consistent disagreement and divergence in estimates.

Driven by these dual objectives, this study conducts a meta-analysis, considering both the temporal variance of reported estimates (Figure 2) and the extent of disagreement in direction and statistical significance (Figure 1). Notably, no meta-analysis has been undertaken, to our knowledge, to probe the actual associations between external interventions and conflict intensity. Our use of meta-analysis methods enables us to amalgamate, synthesize, and probe the estimated parameters, surpassing the confines of an ordinary literature review. A meta-analysis, with its statistical power and regression analysis, surpasses traditional literature reviews by accurately discerning the true direction of empirical findings and unveiling systematic biases such as publication bias. Our overarching aim is to uncover the effectiveness of interventions in conflict mitigation, particularly exploring whether one-sided or general interventions prove more effective, and whether interventions are pro-government or not.

The paper is structured as follows. Section 2 offers a concise review of pertinent empirical literature. Section 3 outlines the steps taken to identify eligible studies, provides an overview of the evidence base, and elaborates on the construction of the meta-dataset. Moving to Section 4, a comprehensive discussion on aspects of estimation design that systematically influence reported estimates of conflict intensity linked to external interventions. Section 5 introduces the methodological approach of the meta-regression analysis (MRA), while Section 6 presents detailed results and discussion. Concluding remarks are

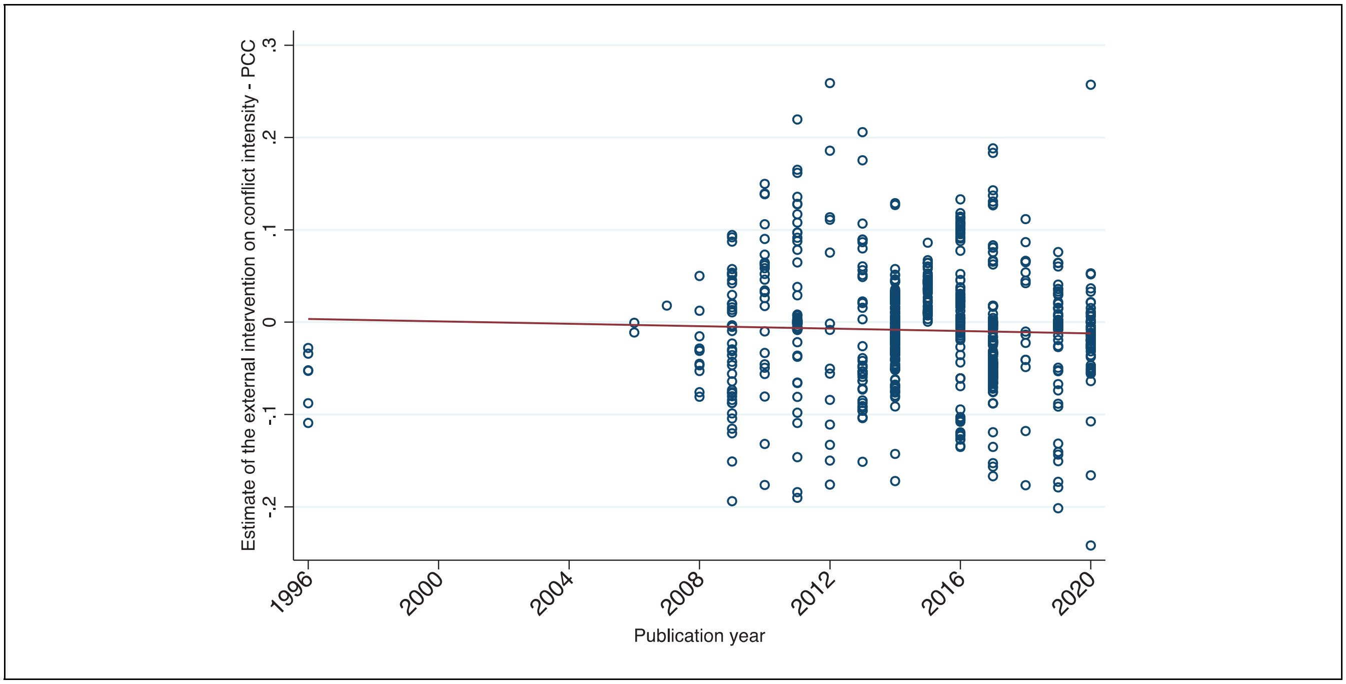

Figure 2. Reported external interventions impacts diverge over the period 1996 to 2020 ($N = 833$).
Note. The figure depicts the impact of external interventions on conflict intensity estimates reported in the individual empirical studies from 1996 and 2020 inclusive. The horizontal axis measures the year when the first drafts of the primary studies appeared in Google Scholar.

provided in Section 7. Additionally, in Appendix A, we present theoretical perspectives on external interventions and conflict intensity are presented.

A Brief Review of Related Empirical Literature

The scholarly exploration of the link between external interventions and conflict intensity is complex, encompassing diverse methodologies that address different facets of intervention. Divergent views exist, with proponents asserting that external interventions can mitigate violence and foster peace, while skeptics question their efficacy. This section offers a concise overview of the key factors centering on the connection between external interventions and conflict intensity.

Numerous studies have explored the consequences of military interventions by external actors in conflict zones. Peacekeeping missions, initially effective in reducing violence and stabilizing situations, often hinge on factors such as mission mandates, troop deployment strategies, and coordination with local actors. Bara and Hultman (2020) assert that the aggregate military capability, encompassing training, equipment, and logistic support to address hostile situations, is a crucial determinant of success. However, the impact of military interventions is not unidimensional. Wood et al. (2012) argue that these interventions may inadvertently escalate violence, particularly when the intervening force is perceived as biased.

By contrast, diplomatic interventions are frequently advocated as alternatives to military involvement. While diplomacy can significantly contribute to conflict resolution, its effectiveness is contingent on factors such as the mediator's credibility, inclusivity of negotiations, and the willingness of conflicting parties to engage constructively. In this context, Regan and Aydin (2006) emphasize the pivotal role of third-party diplomatic interventions in reducing informational asymmetries about capabilities and incentives. External actors facilitate dialog and consensus-building, helping parties find cooperative solutions to end civil wars through negotiated agreements. Other scholars argue that economic assistance to a party involved in a civil war can influence structural conditions that affect the likelihood of conflict prevailing. Regan and Meachum (2014) conclude that economic interventions can increase the chances of regaining stability by addressing political concerns that might otherwise fuel armed rebellion. However, their findings do not indicate a discernible impact from diplomatic interventions.

Identifying the specific types of intervention (economic, military, diplomatic, or peace-keeping missions) employed in the studies included in this assessment is often challenging. However, interveners typically set specific targets that determine whether an intervention is perceived as partisan or neutral. Some studies categorize economic and military interventions as partisan, while diplomatic and mission interventions are often considered neutral. Nevertheless, the lack of explicit specification for the specific target makes retroactively determining the nature of the intervention difficult. For example, interventions involving troops or equipment support, common in military interventions, can be classified as partisan if provided to one of the conflicting parties. Alternatively, they may be viewed as neutral, such as when troops are deployed to monitor a ceasefire or control a buffer zone (Sousa, 2014). Consequently, our empirical approach centers on conflict intensity associated with external interventions, and assesses the specific target (biased or neutral, with biased indicating pro- or anti-government), rather than focusing on the types of interventions.

The empirical literature exploring the relationship between external interventions and conflict intensity reveals a nuanced understanding of the varied outcomes associated with military, diplomatic, and economic interventions. While these interventions have the potential to play constructive roles, their success depends on a multitude of factors. Within scholarly literature, several other potential factors mediate and influence the empirically estimated effectiveness of external interventions. These encompass aspects of the study design used to estimate the relationship, including the choice of data, measures to gauge conflict intensity, the adopted estimation approach, and the specified control variables aimed at explaining the impact of external interventions on the dependent variable — conflict intensity.

Numerous studies examining conflict intensity rely on the monthly count of battle-related deaths as their primary dependent variable. This strategy aims to capture more precise temporal variations and evaluate the specific impact of interventions. For example, Hultman et al. (2014) and Haass and Ansorg (2018) have identified heightened conflict intensity associated with external interventions using this method. Conversely, certain studies, such as those by Di Salvatore (2019), Phayal (2019), and Phayal and Prins (2020), choose yearly counts of battle-related deaths. They argue that this approach offers a more comprehensive understanding, allowing sufficient time to assess the effectiveness of external interventions in reducing violence against civilians. In addition to these approaches, alternative measures are employed in the literature. These include assessing the net impact of interventions, such as the escalation or de-escalation of conflict without explicitly factoring in battle-death counts.

Another segment of the literature focuses on conflict duration, examining the time span from the initiation of a conflict to its resolution or cessation, notwithstanding the fact that precise measurements of civil war duration are notoriously challenging. This body of literature, associated with external interventions, centers on the temporal dimension of conflicts, investigating how long they endure (the length of time) and the factors influencing their duration. In our study, we specifically concentrate on conflict intensity, measured using battle-related deaths. This choice ensures homogeneity and comparability in the reported outcomes. Aggregating diverse outcomes of external interventions would be inappropriate, as emphasized by Demena and van Bergeijk (2017) and Demena et al. (2021).

Nevertheless, numerous studies have delved into the explanatory capacity of conflict duration as a metric of conflict intensity in the context of external intervention. This body of literature suggests that prolonged conflicts may witness heightened intensity, with external interventions influencing these dynamics. However, Kathman and Wood (2014) present somewhat ambiguous findings, indicating that longer wars may encourage more rebel violence in the post-conflict period while potentially reducing militia violence. Several studies also examine the impact of various other control variables on conflict intensity associated with external interventions. For instance, Kathman and Wood (2011) and Di Salvatore (2016) hypothesized that rising income per capita could reduce individuals' willingness to participate in conflict, finding that destitute states are more prone to violent conflict. In a contrasting conclusion, a more recent study by Di Salvatore (2020), found the positive impact of purchasing power on conflict intensity.

Population is another commonly studied control variable. Haass and Ansorg (2018), using a fixed-effect specification, demonstrate that the likelihood of civilian targeting increases in highly populated countries. Additionally, the role of regime type in influencing conflict intensity is explored in various studies. As regimes become more democratic, conflicting parties may rely less on civil victimization, as democracies provide means to address political grievances through peaceful processes (Balcells & Kalyvas, 2014; Kathman & Wood, 2014). Moreover, scholars have investigated the relationship between ethnicity and violence intensity. Studies by Escribà-Folch (2010) and Di Salvatore (2016) indicate that higher levels of ethnic fractionalization are associated with an increased likelihood of civil wars. In diverse societies with multiple ethnic groups, competition and grievances among these groups may contribute to elevated conflict intensity. However, Escribà-Folch (2010) reveals a non-linear relationship between ethnic diversity and conflict intensity.

In our meta-analysis, we scrutinize not only real factors, such as the impact of control variables discussed above, which can shape the explanatory power of external interventions in conflict intensity, but also the influence of study design on estimating this relationship. Researchers commonly utilize time series data to explore the impact of external interventions on conflict intensity, often followed by employing panel datasets. Notably, there is an absence of cross-sectional data application in the literature, potentially minimizing biases related to time-invariant heterogeneity. The analysis of conflict-related data, such as intensity, duration, conflict locations, and casualties, predominantly relies on the Uppsala Conflict Data Project (UCDP).

Regarding study design, the empirical literature varies in the estimation methods employed. The predominant method involves (with or without zero-inflated) negative binomial regression, given the often count-based measure of conflict intensity in terms of battle deaths (e.g., Balcells & Kalyvas, 2014; Beardsley et al., 2019). This method accounts for the over-representation of zeros during active conflict periods without battle-related deaths. Due to the resultant over-dispersed dependent variable, researchers commonly opt for zero-inflated or conventional negative binomial regression. Some other studies employ non-linear regression techniques such as Probit, Logit, or Tobit (see, Dixon, 1996; Doyle & Sambanis, 2000; Kathman & Wood, 2011). A few researchers address endogeneity concerns by applying instrumental variable (IV) techniques to enhance the robustness of causal claims. This recognizes the simultaneous determination of conflict intensity and external interventions, recognizing that intense conflicts may attract more interventions, and vice versa. For instance, Sawyer et al. (2017) use donor GDP as an instrument for external rebel financial assistance in an IV probit model. Similarly, Carnegie and Mikulaschek (2020) employ the rotation of seats between African regions in the UN Security Council as an instrument for the deployment of UN peacekeepers, acknowledging the challenge in identifying precise IVs.

Identifying Eligible Studies and Overview of the Evidence Base

Identifying Eligible Studies

To identify relevant studies, we conducted an extensive search using the Google Scholar web engine, EconLit, and the Web of Science database. Additionally, we performed a manual search using the reference lists of identified recent studies. Employing a broad combination of keywords, such as "external intervention," "external intervention + conflict intensity," "peacekeeping + conflicts," and "third-party intervention + peace process," yielded an extensive literature set. For instance, the combination "external intervention + conflict intensity" yielded 18,400 potential studies. Initial screening based on titles, abstracts, and keywords led to the identification of 98 prospective studies. In cases of uncertainty, further examination was conducted based on the introduction and conclusions of the identified studies.

We selected studies that satisfy the following criteria for detailed review based on the full-text evaluation: English language empirical studies that report econometric results investigating the impact of external interventions on conflict intensity (measured as count of battle death or categorical change of the intensity). The studies must report at least sample size, coefficients, standard errors, or *t*-values. The imposition of these criteria yielded 37 empirical studies published until September 2020 for coding. Majority of the reasons for exclusion were that studies relied on literature review, historical analysis, descriptive studies, other outcome variables, or sanctions as external intervention.

All authors actively participated in constructing the dataset to maintain the highest scientific standard. The two authors independently conducted the search, eligibility assessment, and data extraction using a Microsoft Excel template. Any discrepancies between the two authors were thoroughly reviewed by the third author, and consensus was achieved for a disagreement rate of about 2.5%. Extensive data coding was conducted on various study characteristics, such as conflict intensity, external interventions, data specifications, and publication characteristics. This rigorous process aimed to minimize potential subjectivity bias and enhance the reliability and robustness of the findings.

Overview of the Evidence Base

Figure 3 presents the identified 37 empirical studies, reporting the cumulative number of studies published over the period 1996 to 2020. The evidence started in 1996 by Dixon evaluating the effectiveness of third-party intervention using data for the period 1984 to 1994. The development of the econometric literature was very limited in the 1990s and 2000s. Following this early work, the next study was only published in 2004 examining the impact of various UN interventions for the years 1994 to 1997. Between 2005 and 2010, only a handful of studies (7) was appeared examining the impact of different external interventions (such as peacekeeping, foreign support, non-UN mandate, international mediation) using various cross-country analyses predominantly for sub-Saharan African (SSA) countries. In the last decade (from 2011 to 2020), every additional year resulted in three or four studies. As a result, in that period, a sharp increase in the number of studies was recorded, generating 28 additional empirical studies. The increasing number of studies can be attributed to the expanding availability of data assessing external interventions on conflict intensity, particularly in SSA countries.

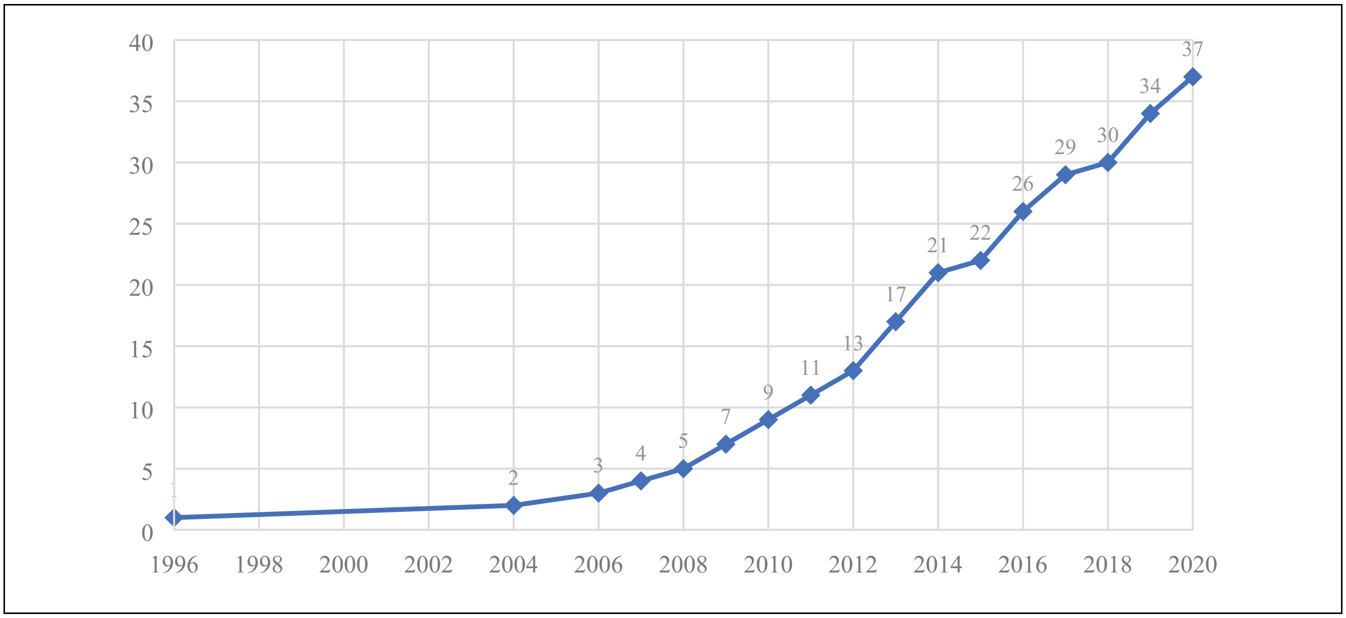

Figure 3. Cumulative number of studies per year published in the period 1996 up to 2020.

It is essential to describe some critical challenges encountered during the coding of reported estimates. A crucial aspect involves deciding between a best-set, average-set or all-set approach in constructing the meta-dataset. Following the recommendation of Stanley and Doucouliagos (2012) and Demena and van Bergeijk (2017), we opted to collect all reported estimates of the primary studies. Another challenge was identifying outliers, with some recent meta-analyses (e.g., Mebratie & van Bergeijk, 2013) considering large estimates exceeding 10 in absolute value as outliers and excluding them from the analysis. In contrast, Stanley and Doucouliagos (2012), suggested unusually large estimates could be attributed to coding errors.

Engaging two independent reviewers, we identified 19 estimates exceeding 10 in absolute value. After meticulous verification, these reported estimates appeared genuine. To impartially evaluate and concurrently pinpoint outliers in the reported coefficients and their estimated precisions, we employed the multivariate method proposed by Hadi (1994). Utilizing this approach, we identified about 4.6% of the reported estimates as outliers, including three studies completely excluded as outliers. This method has been identified as suitable for outliers in multivariate framework and is commonly applied in meta-analysis (Demena, 2021; Demena et al., 2022; Floridi et al., 2020). Applying this method yielded 833 estimates from 34 empirical studies, excluding 40 observations provided by 13 studies. In the Supplemental File, Table B1 presents the list of included studies

Of these studies, 25 are peer-reviewed journal articles, while the remaining nine are working papers or unpublished studies. Over two-thirds of the reported estimates focus solely on evaluating external interventions in Africa, while the rest combining cases from Africa and other regions, except for one study on the Bosnian war. The oldest study was published in 1996, the most recent in 2020 and the median study was published in 2015. This indicates a recent and rapid evolution of the topic, with half of the research published in the last 5 years. The minimum, median and maximum number of observations per study are 1, 18 and 153 coefficients, respectively. For a comprehensive overview, refer to Appendix B, Table B2.

Why Do Estimates of External Interventions on Conflict Intensity Vary? Potential Factors Explaining Heterogeneity

We have observed temporal variation over time in the reported estimates (Figure 2) and substantial differences both within and between studies (Figure 4). This section aims to uncover the underlying reasons for potential sources of heterogeneity across the reported estimates by examining variations in the design of the included empirical studies. We collect various potential moderator variables that reflect each study's data characteristics, estimation characteristics, specification characteristics, and proxies employed for our variable of interests — external interventions, and conflict intensity (Table 1

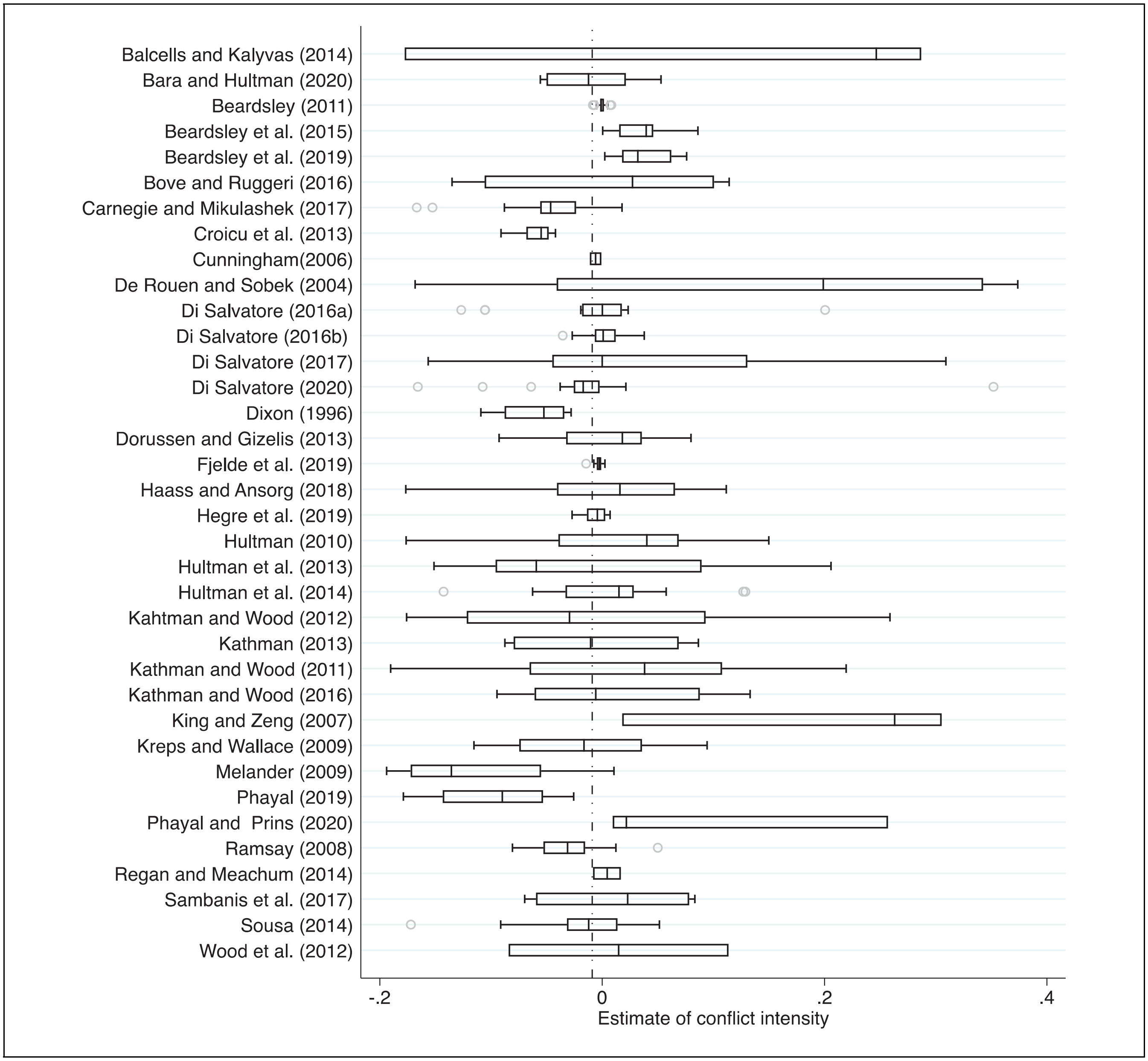

Figure 4. Reported estimates vary widely both within and across the primary studies ($N = 873$).
Note. The figure shows a box plot of the estimates of conflict intensity from external intervention reported in individual primary studies in the period 1996 and 2020. Following Tukey (1977), the box shows the interquartile range (P25–P75), the lower limit the P25 (Q1) and the upper limit the P75 (Q3) with a median represented by a vertical line within the boxes. Horizontal whiskers cover the interval from (P25 − 1.5 times the interquartile range) to (P75 + 1.5 times the interquartile range) (Tukey, 1977). Any dots should show the remaining (outlying) reported estimates in each primary study. The short-dash vertical line represents the arithmetic mean effect of the 873 estimates reported in the primary empirical studies.

gives the details of these selected potential explanatory variables). While we cannot account for all potential moderator variables explaining variations in the reported estimates, our selection is based on common design choices made by researchers of the included studies, aiming to capture the heterogeneity of the reported estimates within the literature.

Data characteristics

The coding of the included studies involves constructing dummies for specific characteristics such as data type (panel or time series), the length or timespan taken by the individual studies, and the number of countries included. Time series data analyses are predominant (82.8%), followed by panel data sets (17.2%), with no application of

Table 1. Description and Summary Statistics of Collected Variables.

Moderator variables	Description	Mean	Std. dev.
Outcome characteristics			
PCC	Partial correlation coefficient	−0.009	0.063
PCC_SE	Standard error of PCC	0.025	0.017
Data characteristics			
Time series	If a time series dataset was used (panel data as reference category)	0.828	0.377
No. years	The logarithm of the number of years of data used	2.945	0.620
No. of obs.	The logarithm of the number of observations	7.863	1.426
Single region	=1 if data come from a single region (e.g., Africa) (data from mixed regions as reference category)	0.667	0.471
Conflict intensity			
Deaths per year	=1 if the outcome variable is a count of battle deaths per year (categorical change in conflict intensity as reference category)	0.176	0.381
Deaths per month	=1 if the outcome variable is a count of battle deaths per month	0.601	0.489
External intervention			
Partisan	=1 if intervention is partisan (neutral or not listed here as reference category)	0.821	0.383
Pro	=1 if intervention is pro-government (neutral or not listed here as reference category)	0.800	0.399
Anti	=1 if intervention is anti-government	0.021	0.141
Estimation characteristics			
Non-linear	=1 if the estimation is multivariate non-linear regression (Probit, Logit, Tobit) (method not listed here as reference category)	0.276	0.447
Negative binomial	=1 if the estimation method is negative binomial regression	0.319	0.466
IV	=1 if the estimation method is 2-stage least square (IV)	0.157	0.364
Specification characteristics			
Natural resource	=1 if the specification controls for the existence of natural resources	0.427	0.494
Wealth	=1 if the specification controls for a measure of wealth or income level	0.541	0.498
Population	=1 if the specification controls for the size of the country's population	0.815	0.388
Year-FE	=1 if the specification controls for year-fixed effects	0.226	0.418
Country-FE	=1 if the specification controls for country fixed effects	0.271	0.445
Polity	=1 if the specification controls for the level of democracy/regime type	0.468	0.449
Ethnic	=1 if the specification controls for ethnic fractionalization	0.155	0.362
Conflict	=1 if the specification controls for the duration of the conflict	0.369	0.483
Publication characteristics			
Publication year	The logarithm of publication year or the age of the study (base, 1996)	2.959	0.298
Reviewed	=1 if published in a peer-reviewed journal	0.471	0.499
Citations	Logarithm of citations in Google Scholar per the age of the study, as of February 2020	1.389	1.217
Impact	The Scimago Journal ranking as of February 2020	44.121	55.271

cross-sectional data in the literature, minimizing the potential biases due to time-invariant heterogeneity (Gujarati & Porter, 2009). The mean logarithm values of the number of years of data and the observations per study are 2.945 and 7.863, respectively. About 66.7% of estimates originate from studies analyzing the impact of external inventions on conflict intensities in a single region (Africa, Europe or Asia). This is important as regional differences may influence the effectiveness of external interventions, considering that some regions experience more conflicts than others. For instance, Africa, with 67% of reported estimates, has encountered more conflicts and consequently attracts a greater number of interventions. In 2005 alone, 77% of UN peacekeeping operations were deployed in Africa, constituting 75% of the total budget of that year (Murshed, 2009).

Dependent variable (conflict intensity)

To capture a better temporal variation in conflict intensity as well as to provide a better effect of an intervention, most studies deploy the monthly counts of battel-related deaths as their dependent variable. As a result, about two-thirds (60.1%) of the reported estimates build from a monthly account of battel-related deaths. The second dependent variable is measured in terms of the yearly count of battel-related deaths. These studies (for instance, Di Salvatore, 2019; Phayal, 2019; Phayal & Prins, 2020) argue that measuring conflict intensities over the year yields a better understanding than monthly as it takes time to assess external interventions to curtail violence against civilians. Whereas the remaining estimates had used other measures such as the net impact of interventions (e.g., escalation of conflict, de-escalation of conflict or no change in battle deaths). Our analysis centers on examining the impact of external interventions or third-party involvements on conflict intensity measured in terms of battel-related death rather than assessing the success or failure of these interventions.

Independent variable (external interventions)

The international community's interventions often categorized as economic, military, diplomatic or a combination, are implemented by States or international organizations like the UN on conflict-affected States (Sousa, 2014). While identifying the specific intervention types (i.e., if economic, military, diplomatic or mission) in the included studies may not often be possible, interveners generally have specific targets, such as supporting opposition or government or maintaining neutrality. We therefore first classify interventions as partisan/biased or neutral based on the studies, with biased interventions supporting either the government or the opposition. We considered that third-party intervention is partisan/biased, when supporting the opposition or the government, and neutral, when the target is not intended to change the balance of capabilities of the parties. In over 82.1% of estimates, interventions were considered partisan. Next, we classify, whether the type of partisan intervention is in supporting the government (current political system) or opposing group or analyzing neutral mediations. We found that the majority (80.0%) of the reported estimates from the primary studies involve intervention as pro-government, while the remaining assess anti-government or neutral mediations.

Estimation characteristics

Researchers in this literature adopted various estimation techniques to examine the relationship between external interventions and conflict intensity nexus. The most applied method is negative binomial regression (31.9%), with or without zero-inflation adjustment. Since conflict intensity is mostly measured in counts of battel deaths, there is a common issue of over-representation of zeros and over-dispersed dependent variable (Phayal & Prins, 2020). This occurs when months or years have active conflict but without any battle-related deaths, resulting in a choice of dependent variable coded as 0 battle deaths (Beardsley et al., 2019). In this case, the proportion of monthly or yearly counts of conflict without battel-related deaths is substantially higher than for the cases with nonzero battle deaths. Similarly, it is normal to observe that the variance of the dependent variable is greater than its mean, suggesting the outcome variable is highly over-dispersed. Given this nature, either zero-inflated regression or a conventional negative binominal regression is the preferred estimation approach. Another quarter of the reported estimates (27.6%) in primary studies employed non-linear regression techniques such as Probit, Logit, or Tobit. Approximately 15.7% of estimates addressed endogeneity concerns by using instrumental variable (IV) techniques, like two-stage least squares, as discussed in Section 2.

Specification characteristics

In terms of model *specifications*, various control variables have been considered. Dummies were included to account for factors such as natural resources, income level, country's population, level of democracy or regime type, ethnic fractionalization, and conflict duration. Recent empirical studies have highlighted the significance of these variables in assessing the impact of external interventions on conflict intensity. Natural resources (Carnegie & Mikulaschek, 2020; Di Salvatore, 2020), income level (Kathman & Wood, 2014), country population (Bara & Hultman, 2020), democracy or regime type (Kathman & Wood, 2011), ethnic fractionalization (Balcells & Kalyvas, 2014), and conflict duration (Bara & Hultman, 2020; Beardsley et al., 2019; Haass & Ansorg, 2018) are among the key considerations. Additionally, year- and country-fixed effects accounted for 22.6% and 27.1% of the estimates, respectively. Fixed effects specification is preferred for country-level variations, providing better handling of temporal factors and other poorly measured static variables (Phayal & Prins, 2020).

Publication characteristics

We also accounted for the *publication characteristics* of the included studies. We attempt to control for potential aspects of the quality of the included studies looking at factors that are not reflected by the heterogeneity in empirical designs and data characteristics discussed above. We control for study quality to test if peer-reviewed publication is systemically linked with reported external intervention estimates. About half of the reported estimates are contributed by publications in peer-reviewed journals. Additional aspects of study quality considered encompass the number of citations in Google Scholar relative to the age of the study, the publication year of the study and the Scimago Journal ranking.

Meta-Regression Methodology

A standardized methodological approach is essential when dealing with estimates that vary in proxies for external interventions and conflict intensity outcomes. To ensure comparability across included empirical studies, we employ the partial correlation coefficient (PCC), a widely applied method in recent meta-analyses (Demena, 2021; Floridi et al., 2020, 2022). The PCC is computed as:

$$PCC_{is} = \frac{t_{is}}{\sqrt{t_{is}^2 + df_{is}}} \quad (1)$$

where PCC_{is} denotes the partial correlation coefficient between external interventions and conflict intensity. This measures the association in terms of the direction and strength of these two variables holding other variables constant. t_{is} signifies the corresponding t-values and df are the associated degrees of freedom for the reported estimates in each of the regression specifications of the studies. $s = 1, \ldots, 34$ represents the primary study; $i = 1, \ldots, 833$ represents the reported estimates specified in each of the regression of the primary studies.

After deriving the PCC, our approach follows three stages. In the first stage, we compute both the simple overall average effect of external interventions and the weighted effect using inverse variance weights. In the second stage, we employ both graphical and statistical analyses. We begin with a graphical assessment to determine the authenticity of the overall effect derived in the first stage and to identify any potential selection bias. Funnel plots are employed to get an indication of the extent of publication bias. While this method relies on visual inspection, which can be subjective, it is complemented with formal statistical analyses.

To formally filter out potential publication bias and derive the genuine effect, we apply the following meta-regression model (MRM):

$$PCC_{is} = \beta_0 + \beta_1 SE_{pccis} + u_{is} \quad (2)$$

where PCC_{is} is is the measure of estimated external interventions effect for the regression specification i reported by the study s. SE_{pccis} is the associated standard error of the PCC (computed as $\sqrt{(1 - {pcc^2}_{is})/ df_{is}}$). The coefficient β_1 is the publication bias and β_0 is the estimated genuine overall effect after controlling for potential publication bias. Equation 2 indicates that as the quantity of available information increases, that is, as N increases, standard error, SE_{pccis} will approach nil. That means, with a large sample size for a given regression specification, PCC_{is} will approach the overall size of the genuine effect, β_0, irrespective of publication bias. The presence of publication bias, however, can be detected if PCC_{is} is associated with its respective standard error, SE_{pccis}.

In a classical regression, the error term (u_{is}) must be independent and identically distributed to obtain an efficient estimator (Demena & Afesorgbor, 2020). In estimating Equation 2 however, the SE_{pccis}, which is the independent variable, is the standard error of the dependent variable, PCC_{is}. As a result PCC_{is} has a different estimated variance, and thus empirically involves the problem of heteroscedasticity. This is also indicated as the problem of the likelihood of intra-study dependence in error terms (Stanley, 2005). To reduce this problem, as well as to assign greater weight to estimates with greater precision, Equation 2 should be estimated with a weighted least squares (WLS) approach. This follows that Equation 2 is weighted by inverse variance of the estimated PCC_{is}, which is dividing by SE_{pccis} and becomes:

$$t_{is} = \beta_1 + \beta_0(1/SE_{pccis}) + e_{is} \quad (3)$$

where t_{is} is the t-value measuring the statistical significance of the PCC computed from PCC_{is}/SE_{pccis}. Using Equation 3, we then investigated the presence of publication bias by testing the hypotheses that $\beta_1 = 0$. This is known as the funnel asymmetry test (FAT). Similarly, we explore the presence of the overall underlying effect of external interventions by testing the statistical significance of β_0, the coefficient associated with the inverse of the standard error, $((1/SE_{pccis}).)$. The parameter β_0 therefore estimates both the magnitude and direction of the underlying genuine effect, known as the precision-effect test (PET).

In the third stage, we incorporate the sources of heterogeneity outlined in Section 4 to elucidate the factors contributing to the diverse reported estimates, as evident in Figure 4, exhibiting substantial variations in reported estimates within and across primary studies. In this stage, we endeavor to explore the extent to which the estimates reported in each study are influenced by the nature of the empirical design and other study characteristics. To address heterogeneity, we enhance the FAT-PET in Equation 3, leading to the setup of a multivariate MRM as described in Equation 4.

$$t_{is} = \beta_1 + \beta_0(1/SE_{pccis}) + \alpha_k X_{kis}/SE_{pccis} + e_{is} \quad (4)$$

All other parameters remain consistent with those reported in Equation 3. Here X denotes a vector encompassing the moderator variables outlined in Table 1 along with their summary statistics and descriptions. α_k represents the corresponding coefficient, with K representing the specific category of the moderator variables.

Beyond within-study dependence, estimating both Equations 3 and 4 involves additional concerns. An important concern is the potential for between-study dependence, especially when multiple studies are authored by the same individuals, as in our case, rendering them statistically non-independent. To address this, we apply both the basic weighted clustered standard errors (CDA) model, and the mixed-effects multilevel (MEM) model (Demena 2024a). For the CDA, we mean study-level clustered standard errors, hence the reference to clustered data analysis. We also employ a fixed effect (FE) estimation that may also address the issue of individual within-variation (Demena & Afesorgbor, 2020). Additionally, one may argue that the number of primary studies seems to be limited, and thus we also employ the wild bootstrap clustering, tailored for cases with a small number of clusters, considering the perceived limitation in the number of primary studies (Demena, 2024b). For the MEM, we initially use Cochran's Q-test to assess the presence of statistical dependency between studies, as highlighted in Table 3.

To capture the statistical dependency between studies while also controlling for within-study dependence, a multi-level model or hierarchical model is considered as most appropriate (Demena, 2015). Through the inclusion of a random individual effect for each study, this model allows to account both the within-study and between-study dependences and thus the reference to a multilevel random effect model (Demena, 2021). The latter is best designated as MEM controlling for both within-study and across-study variations when it includes the identified moderator variables (Afesorgbor et al., 2024; Demena, 2017). Accordingly, this is described as a two-level model, augmenting Equation 4 becomes:

$$t_{sc} = \beta_0 + \beta_1(1/SE_{pccsc}) + \frac{\alpha_k X_{ksc}}{SE_{pccsc}} + \zeta_s + e_{sc} \quad (5)$$

ζ_s represents the study-level random effects (random intercepts). All other parameters remain consistent with those outlined in Equation 4. In this notion, the reported estimates, which are level 1, are clustered and nested within studies, which are level 2. Consequently, the estimates pertaining to external interventions reported in each empirical study are nested within that particular study and the model allows for variations between studies.

Another empirical concern of estimating Equation 4 or 5 is the potential multicollinearity which can arise if many moderators are employed as presented in Table 1. Adhering to the MAER-Net guidelines (Havránek et al., 2020) and in line with contemporary trends in meta-analysis (Demena & Afesorgbor, 2020; Floridi et al., 2021, 2022), we employ the general-to-specific (G-to-S) approach. The procedure commences by including all potential moderator variables in the general specification and subsequently systematically eliminating the least significant variables one at a time until only the significant ones remain. Doing so, we exclude 13 variables that are statistically insignificant at least at the 10% level. The joint-test for the remaining nine moderator variables rejects the null hypothesis of a zero joint effect, (F_9, 808) = 8.92 (p = 0.000). Conversely, in support of the excluded 13 moderator variables, the null hypothesis of a zero joint effect cannot be rejected, (F_{13}, 808) = 0.52 (p = 0.9163). The latter indicates that the 13 moderator variables jointly appear to be statistically indistinguishable from zero, and thus do not contribute to the explanation of the heterogeneity.

Table 2. Overall Reported Estimates of External Intervention on Conflict Intensity.

Method	Effect size	S. E.	95% Confidence interval	
Simple average effect[a]	−0.009	0.002	−0.013	−0.005
Weighted average effect[b]	−0.007	0.002	−0.010	−0.004

[a]The arithmetic mean of the conflict intensity impact associated with external interventions.
[b]Uses inverse variance as weight.

Meta-Regression Results

Bivariate Meta-Regression Evidence

Prior to conducting the bivariate MRA, we compute both the simple and weighted average effects. Table 2 shows the overall impact of external interventions on conflict intensity.

The unweighted simple average across studies reveals a statistically significant negative effect (−0.009) of external interventions on conflict intensity. To enhance precision, we also compute the weighted average effect using inverse variance. Despite a 22% reduction in magnitude, the weighted average effect aligns with the unweighted average, signifying a negative and statistical impact. The evidence, therefore, suggests that external interventions are likely to significantly diminish conflict intensity. However, following established guidelines for PCC effect magnitude (Doucouliagos, 2011), the practical impact of external interventions on reducing conflict intensity appears nearly negligible. While this initial summary provides insights, it serves as an indication. To offer a more comprehensive understanding, we must address potential publication bias and explore various drivers of heterogeneity. Accordingly, our next step involves addressing issues related to publication bias.

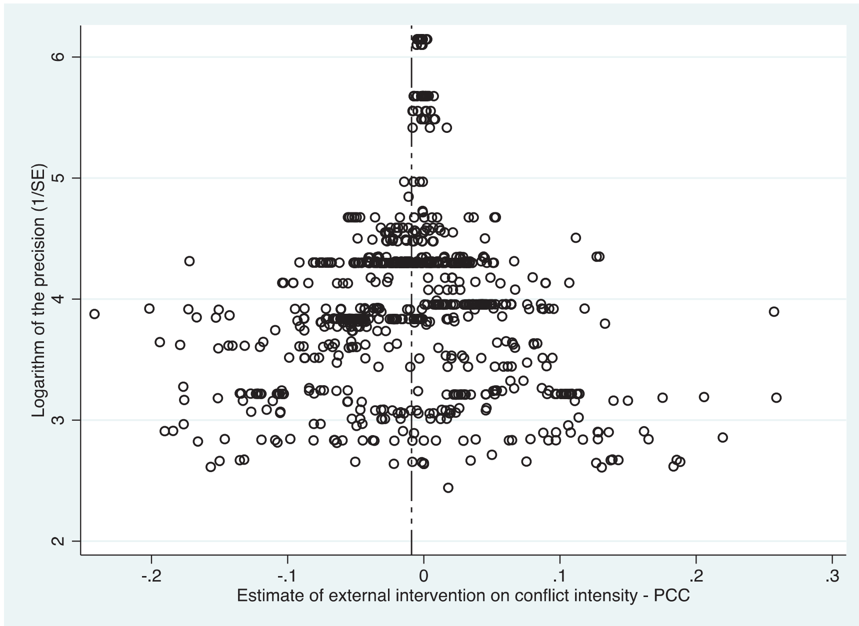

Figure 5. Funnel plot testing publication bias ($N = 833$).
Note. To allow better visualization of the plots between reported effect size and their precision, we present the logarithm of the inverse of the standard error. The long-dash short-dash vertical line represents the inverse variance weighted average effect of the full sample as reported in Table 2.

Testing for Publication Bias: Graphical Inspection. Funnel plot is a widely method to assess the issue of publication bias. It takes the form of a scatter diagram, where reported external interventions are plotted on the horizontal axis and their precision (typically the reciprocal of the standard error) on the vertical axis. In the absence of publication bias, a funnel plot exhibits symmetry, resembling an inverted funnel (Stanley, 2008). This symmetry arises because the most precise estimates are clustered around the underlying effect at the top of the funnel, while less precise estimates are more widely dispersed at the bottom of the graph (Stanley, 2005). Conversely, the presence of publication bias becomes apparent when examining an asymmetrical funnel plot. In this case, a tendency exists to selectively report statistically significant results in a given direction. This implies that certain estimates are either overrepresented or discarded to favor a particular expected sign or conclusion (Afesorgbor & Demena, 2022; Demena, 2021). There is also the potential publication bias in a seemingly symmetrical funnel plot when the graph appears to be hollowed and excessively wide. As highlighted by Stanley (2008), this occurs when statistically significant results are chosen irrespective of the direction of the effects, or when both statistically significant negative and positive findings receive equal reporting preference, resulting in the appearance of a symmetrical funnel.

Figure 5 gives the funnel plot. The vertical axis displays the logarithm of the inverse of the standard error and the horizontal axis the magnitude of the estimated external intervention effect. The plots appear to have a symmetrical funnel plot. The most precise estimates are close to around the weighted effect (−0.007), represented by the long and short dashed vertical line. Many imprecise reported estimates seem to be evenly distributed on both the right and left sides of the overall weighted average effect. In other words, there are estimated effects both larger and smaller than −0.007. This implies that the reported estimates of the impact of the external interventions exhibit a symmetrical distribution around the vertical axis (weighted average effect), without any noticeable hollowing or excessive width in the reported estimates. Therefore, the funnel plot suggests an absence of any publication selection bias. However, it's important to note that this method of testing publication bias is based on visual inspection, making it subjective and less compelling. Hence, the application of the formal statistical method of MRA is required.

Testing for Publication Bias: Statistical Analysis. Moving beyond the basic visual inspection of the funnel plot, we employ econometric testing to ascertain the presence of publication bias. Doing so, we introduce a more formal

Table 3. Publication Bias and Underlying Genuine Effect Test.

All studies—FAT/PET	(1)	(2)	(3)	(4)
Variables	CDA	FE	MEM	Wild bootstrapped
Publication bias (FAT—constant)	−0.459	−0.508	−0.380	−0.459
	(0.399)	(0.102)	(0.285)	0.368
Genuine effect (PET—precision)	−0.007	−0.0004	−0.0003	−0.001
	(0.002)	(0.001)	(0.002)	0.67
Observations	833	833	833	833
Studies	34	34	34	34
LR test			132.54	
$p > \text{chi}^2$			0.000	

Note. Reported results are estimated from Equation 3 specification. All estimates use the inverse variance as weights and standard errors reported in parentheses are clustered at the study level. Panel 1 (CDA: clustered data analysis is estimated via the study level clustered robust standard errors; Panel 2 (FE) is fixed-effect estimation clustered at the study level; Panel 3 (MEM) is mixed-effects multilevel estimated through the restricted maximum likelihood; and Panel 4 (Wild bootstrapped) is regression bootstrapping the standard error (a non-standard cluster adjustment) reported with *p*-values. Test for between-study heterogeneity (Cochran's *Q*-test) is 55521.34*** on 832 degrees of freedom with *p*-value less than .001 and I^2statistics (variation in reported estimates attributable to heterogeneity) is 84.9%.

method for testing bias through regression-based FAT, as detailed in Section 5, Equation 3.

Table 3 reports the results of the FAT analysis using various estimation specifications. Across all estimations, the FAT does not indicate the existence of a statistically significant publication selection bias, despite all estimated coefficients being negative. Across the specified estimation techniques, publication bias ranges between −0.380 and −0.508. Following the perspective of Doucouliagos and Stanley (2013), irrespective of the applied econometric specification, the magnitude of the estimated publication bias in the research of external interventions and conflict intensity is considered irrelevant. Therefore, in alignment with the visual inspection of the funnel plot, the formal statical approach (FAT) consistently reports a clear absence of evidence for bias.

The Underlying Genuine Effect. Moving beyond the issue of publication selection bias, Table 3 also allows an examination of the presence of a genuine underlying effect. We apply the method of PET specified in Equation 3 that provides the overall average effect after accounting for publication selection bias. Consistent with the average effects presented in Table 2, all PET results across the estimation methods consistently point to a negative impact of external interventions, but statistically insignificant. This implies that external interventions are not associated with a significant reduction in conflict intensity.

Nonetheless, it is worth indicating that these findings represent average across all empirical methods. As presented in Figure 4 (depicted in the box plots), the reported estimates in the studies exhibit substantial variability both within and between studies. Empirically, the between studies heterogeneity is evident by Cochran's *Q*-test, as reported in Table 3. Therefore, a multivariate MRA, as presented in the next section, becomes crucial, as our inferences may also depend on various potential sources of heterogeneity, including the quality of the primary studies, misspecification, research design or other characteristics. Moreover, this approach will further refine the examination of publication bias and the overall genuine underlying effect.

Multivariate Meta-Regression Evidence: Explaining Observed Sources of Heterogeneity

Table 4 gives the results of the reduced multivariate MRA estimating Equations 4 and 5 using the G-to-S modeling. Following this approach, outlined in Section 5, Column 1 presents the specific multivariate model with 9 variables that remain statistically significant, at least at the 10% level. To address within-study correlation, this specific model (Column 1) is subsequently re-estimated by applying the preferred MEM model (Column 4, Equation 5). For comparison and robustness checks, we also include secondary specifications (Equation 4) of CDA in Column 2, FE in Column 3, and wild bootstrap clustering in Column 6. Additionally, in Column 5 we replace partisan external intervention with pro-government and anti-government interventions for further analysis.

Based on these estimations, we identify seven points highlighting observed sources of heterogeneity and variations in the underlying meta-effect. Conclusions are drawn primarily from our preferred specification

Table 4. Explaining the Drivers of Heterogeneity in the Estimates Reported Across Studies.

	(1)	(2)	(3)	(4)	(5)	(6)
Moderator variables	Specific	CDA	FE	MEM	MEM	Wild bootstrapped
Bias coefficient	0.675***	0.675**	−1.119**	0.257	0.102	0.675**
(Constant-FAT)	(0.191)	(0.272)	(0.457)	(0.370)	(0.379)	0.01
Genuine effect	−0.056***	−0.056***	−0.098	−0.056*	−0.048	−0.056**
(Precision-PET)	(0.019)	(0.020)	(0.168)	(0.032)	(0.033)	0.002
Data characteristics						
Single region	−0.016***	−0.016***	−0.015	−0.012*	−0.010	−0.016**
	(0.004)	(0.005)	(0.024)	(0.007)	(0.007)	0.018
Conflict intensity & external intervention						
Death per year	−0.040***	−0.040**	−0.058	−0.035**	−0.037**	−0.040**
	(0.009)	(0.015)	(0.071)	(0.017)	(0.017)	0.024
Partisan	−0.015***	−0.019*	−0.014	−0.013**		−0.015
	(0.005)	(0.011)	(0.016)	(0.005)		0.472
Pro Gov.					−0.014**	
					(0.005)	
Anti Gov.					0.076***	
					(0.027)	
Estimation characteristics						
IV	−0.040***	−0.040***	0.081***	0.011	0.015	−0.040**
	(0.006)	(0.006)	(0.003)	(0.019)	(0.019)	0.002
Specification characteristics						
Year FE	−0.016***	−0.016*	0.005***	−0.008	−0.009	−0.016
	(0.005)	(0.009)	(0.001)	(0.007)	(0.007)	0.192
Natural resources	−0.011**	−0.011*	0.005	−0.002	−0.002	−0.011*
	(0.004)	(0.005)	(0.005)	(0.007)	(0.007)	0.100
Publication characteristics						
Publication year	0.027***	0.027***	0.040	0.023**	0.020*	0.027***
	(0.006)	(0.007)	(0.060)	(0.011)	(0.011)	0.000
Published	−0.027***	−0.027**	0.009	−0.010	−0.005	−0.027
	(0.007)	(0.013)	(0.057)	(0.014)	(0.014)	0.134
Journal impact	0.000***	0.000	−0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	0.266
Observations	833	833	833	833	833	833
Studies	34	34	34	34	34	34
LR test				−1888.61	−1882.95	
$p > \text{chi}^2$				*0.089*	*0.004*	

Note. *, **, and *** stand for 10%, 5%, and 1% level of significance, respectively. Reported results are estimated from Equations 4 and 5 specifications. All estimates use the inverse variance as weights and standard errors reported in parentheses are clustered at the study level. Panel 1 is the specific model derived from the general specification (G-to-S approach); Panel 2 (CDA: clustered data analysis, also known as WLS) is estimated via the study level clustered robust standard errors; Panel 3 (FE) is fixed-effect estimation clustered at the study level; Panel 4 (MEM) is mixed-effects multilevel estimated through the restricted maximum likelihood; Panel 5 (MEM) replaced the external intervention partisan with pro-government and anti-government interventions; and Panel 6 (Wild bootstrapped) is regression bootstrapping the standard error (a non-standard cluster adjustment) reported with p-values.

(Column 4) and are categorized as follows: (i) highly consistent evidence if both sign and significance are supported by two secondary specifications (Columns 2 or 3, and 6); (ii) moderately consistent evidence if consistency exists with one of the secondary specifications (Column 2, 3, or 6); and (iii) weakly consistent if consistency is observed between two secondary specifications only (among Columns 2, 3, and 6) or results are supported by preferred specification only (Column 4).

First, we find highly consistent evidence that a single region examination tends to yield lower conflict intensity compared to a multi-region analysis. Our findings suggest that higher conflict intensity in primary studies that include countries from different regions (cross-region analysis) is likely to be upward biased due to aggregation bias in the aggregated data. bibr3 Therefore, we can infer that a single region analysis may be a better approach or good practice in future research, as it permits capturing facts more closely and identifying a more precise impact of external interventions.

Second, we find highly consistent evidence of systematic difference in terms of the outcome variable used to measure the extent of conflict intensity. Specifically, compared to other measures of conflict intensity, the impact of external intervention yields lower estimates when deaths per year is used as a proxy variable. Our results align with findings in the literature (e.g., Di Salvatore, 2019; Phayal, 2019; Phayal & Prins, 2020). These studies suggest that measuring conflict intensities over the year provides a better understanding than other measures, such as battle deaths per month. They argue that it takes time to assess the current and future potential for peacekeeping missions to curtail violence against civilians. Therefore, in future research, analyzing annual data would offer clearer insights into how interventions or observer missions are effective.

Third, concerning external interventions, we find moderately consistent evidence indicating that the type or target of intervention systematically affects the extent of conflict intensity. Specifically, we find that partisan intervention, as opposed to neutral intervention, is associated with de-escalation of the conflict. This aligns with the existing literature, which posits that neutral interventions tend to escalate conflict, while partisan interventions enhance the chance of victory for the supported party, thus being linked to reducing battlefield deaths (Hultman et al., 2014). Conversely, Sousa (2014) finds that partisan interventions escalate conflict intensity, but neutral interventions do not have any effect.

To provide a more nuanced assessment, we further explore whether partisan interventions is supporting the government (pro-government) or opposition group (anti-government). We find that if third-party intervention activities are pro-government, they significantly reduce battle deaths, whereas anti-government interventions enhance such deaths. This suggests that interventions supporting the government tend to have adequate funding and military deployment, while those supporting opposition sides do lend to "cheap talk" outcomes. These results confirm the previous findings in the literature, for instance, Haass and Ansorg (2018) and Sambanis et al. (2020), who observed that governments get support through peacekeeping operations as these foreign troops are more experienced, equipped with high-quality militaries, and are better able to deter violence from state and non-state actors. Additionally, these missions can create buffer zones within conflict areas, reach remote locations more effectively, and exert diplomatic pressure through peace agreements (Haass & Ansorg, 2018). Conversely, external interventions in the absence of popular local support may act as a catalyst for civil war, particularly when combined with other factors like ethnic or social fractionalization (Sambanis et al., 2020). These results have significant implications, suggesting that when partisan intervention supports the government, it is likely to de-escalate conflict intensity, while the opposite holds true for one-sided interventions in favor of rebels, in terms of battle deaths.

Fourth, regarding estimation characteristics, the application of IV stands out, as compared to other estimators, by yielding consistently negative and significant estimates in most cases (except for MEM models). Recent meta-analyses (e.g., Anwar & Afesorgbor, 2021) reported that IV specifications, controlling for endogeneity, measurement bias, unobserved country-fixed effects, and other potentially omitted variables are likely to yield more precise estimates in line with the theoretical predictions. This importance of IV aligns with those of Sawyer et al. (2017), who, using donor GDP as an instrument for external rebel financial assistance, discovered that rebels receiving this highly fungible support are less likely to see the end of the conflict. Additionally, our results show that studies controlling for year-fixed effects, on average, yield lower estimates than those that do not. While several recent meta-analyses suggest that dealing with unobservable time-varying effects can reduce potential bias in impact analysis, the results of these associations are weakly consistent across different specifications.

Fifth, the inclusion of natural resources in the estimated models presents a noteworthy consideration. We find weakly consistent systematic differences between studies that include the existence of natural resources in the empirical models and those that do not. This result aligns with existing work related to "greed and grievance" (e.g., Sambanis et al., 2020; Sousa, 2014). However, contrary findings have been documented by de Rouen and Sobek (2004) and Di Salvatore (2020).

Sixth, regarding publication characteristics, we find highly consistent evidence that the publication year (age of the study) systematically affects conflict intensity estimates. Recent studies tend to report an enhancing role of external interventions in conflict intensity, supporting the "research-cycle" hypothesis regarding novelty and fashion in academic research (Goldfarb, 1995). This hypothesis proposes seminal research often produces promising and significant impacts, while skeptical evidence follow-up with contrary findings of the empirical effect in question (Afesorgbor & Demena, 2022; van Bergeijk et al., 2019). Moreover, we find no systematic difference between estimates reported in peer-reviewed articles and working papers, as well as the journal's impact factor. This finding suggests that publications in peer-reviewed journals (as opposed to working papers) do not tend to report selected evidence. This contrasts with the argument in Costa-Font et al. (2013), known as the "winner's curse," suggesting that peer-reviewed articles are associated with more selected evidence.

Finally, having considered these observed sources of variation, we offer a holistic view of the underlying implied effect of external interventions on conflict intensity. In comparison to the bivariate analyses, the inclusion of the observed moderator variables strongly impacts the significance and magnitude of the PET but has a weak impact on FAT. Regarding FAT, we observe weakly consistent publication bias with a magnitude of little selectivity, indicating the bias found in this literature is irrelevant. Concerning PET, there is now an underlying genuine effect between external interventions and conflict intensity. However, there are many underlying genuine heterogeneity effects than that are only related to a single PET (Demena & van Bergeijk, 2017). Therefore, we apply the "best practice" underlying overall genuine effect, exploring the findings presented in Table 4 (for a similar approach, see Afesorgbor & Demena, 2022; Floridi et al., 2020). We define the best practice conditional on a set of moderator variables, consisting of coefficient estimates derived from recent studies that focus on a single-region empirical model, analyze the outcome indicator as yearly battle-related deaths associated with partisan interventions (support of government), and the design being an IV approach that includes time-fixed effects.

The resulting best practice effect predicts a genuine effect conditional on the identified heterogeneity, which is −0.090 and statistically significant at the 5% level. Therefore, we observe that the corrected correlation coefficient derived from the best practice approach is substantially larger than the simple (weighted) average as well as the bivariate PET. Importantly, it is statistically significant after accounting for the observed sources of heterogeneity. However, following Doucouliagos (2011), the implied genuine overall effect corresponds to a small-sized correlation coefficient. Hence, by accounting for potential observed moderator variables and applying the best practice approach, the current evidence suggests that external intervention efforts, overall, prove advantageous in curbing or de-escalating conflict intensity. It is noteworthy, however, that the magnitude of the statistical significance coefficient is small. For policy purposes, external interventions are likely to have a modest effect. Any policy intervention involves some form of cost-benefit analysis, involving in this case, not just the scale of the desired outcome of the intervention but also the implementation costs linked to the intervention.

Conclusions

We present the first meta-analysis encompassing a comprehensive array of empirical studies spanning the 1996 to 2020 period. Our analysis delves into the impact of external interventions on conflict intensity, drawing on a database comprising 833 reported estimates from 34 distinct studies. While the aggregated reported effect emerges as negative and statistically significant, a noteworthy degree of variability characterizes these estimates. Approximately one-third of the reported estimates indicate a negative and statistically significant effect, offering a clear pattern. In stark contrast, a positive and statistically significant impact is observed in roughly one-fifth of the reported estimates. Meanwhile, the remaining 44% of estimates reveal evidence of statistical insignificance, whether negative or positive. Our meta-analysis aims to disentangle the diverse sources contributing to this variation, systematically examining the factors that exert systematic effects on the reported estimates. Furthermore, our work involves systematic exploration of the overall genuine effect underscored by the collective findings of this literature. Adhering strictly to the MRA reporting guidelines of MAER-Net, we present a nuanced and thorough analysis that sheds light on the intricate dynamics at work within the realm of external interventions and conflict intensity.

First, our analysis reveals a dependence of the reported estimations on the type of data utilized. Single-region data tends to yield lower estimations of conflict intensity compared to samples incorporating mixed/multi-region data. This suggests a potential upward bias in samples with aggregated multi-region data due to aggregation bias. Consequently, we advocate for a shift toward a single-region analysis in future research, positing it as a more prudent approach. This approach allows for a closer examination of the impact of external interventions.

Second, our findings indicate that estimation outcomes are influenced by both the target of interventions and the measure of conflict intensity. Notably, partisan interventions exhibit a trend toward de-escalation, contrasting with neutral interventions that seem to escalate conflict. This pattern suggests that partisan interventions contribute to the increased likelihood of victory for the supported party, resulting in a reduction of battle-related deaths. Additionally, pro-government third-party interventions demonstrate a capacity to diminish battle deaths, while the converse holds true for anti-government interventions. This implies that interventions supporting governments tend to benefit from ample funding and military deployment, while those supporting opposition factions may lend to the "cheap talk" interventions. Furthermore, our analysis indicates that the impact of external intervention on conflict intensity is more nuanced when annual battle deaths are considered, as opposed to using monthly data, which doesn't systematically influence reported outcome significance. This suggests that the temporal dimension, specifically the time required for effects to materialize, plays a crucial role in mitigating conflict intensity. As a recommendation for future research, employing annual conflict intensity data is advised to gain clearer insights into the effectiveness of interventions or observers' missions.

Third, our investigation uncovers a noteworthy influence of publication year on reported estimates, with recent studies indicating an amplifying role of external intervention in conflict intensity. This phenomenon may be attributed to the "research-cycle" hypothesis, wherein novel and trendy academic research tends to yield promising and significant impacts, followed by more skeptical evidence in subsequent studies. Alternatively, it might suggest a transformation over time, where initially well-funded and serious external interventions gradually devolve into mere "cheap talk." While our findings do not provide compelling evidence, we do observe either weakly or moderately systematic variations related to estimation techniques, time-fixed effects, and the presence of natural resources.

Considering these various observed sources of variation, we present the overall implied effect of external intervention on conflict intensity, conditioned on the identified heterogeneity. The resulting best-practice effect predicts a genuine effect of -0.090, which is statistically significant at the 5% level. This suggests that external intervention activities make a statistically significant contribution to restraining or de-escalating conflict intensity, albeit with a practical effect of small magnitude. For policy purposes, therefore, external interventions are likely to have a modest effect. Any policy intervention involves some form of cost-benefit analysis, involving in this case, not just the scale of the desired outcome of the intervention but also the implementation costs linked to the intervention. Importantly, our analysis does not reveal any systematic evidence of publication selection bias in the literature on external interventions and conflict intensity.

Acknowledgements

A preliminary version of this paper was presented at the 20th Jan Tinbergen European Peace Science Conference (2021) and at the Institute of Development and Policy Research, Addis Ababa University. Comments by participants are gratefully acknowledged. The author acknowledges the financial assistance of the Development Economics research group, International Institute of Social Studies, Erasmus University.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Binyam Afewerk Demena https://orcid.org/0000-0002-9059-5364
Amar Anwar https://orcid.org/0000-0002-9247-8452
Syed Mansoob Murshed https://orcid.org/0000-0002-8360-033X

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Supplemental Material

Supplemental material for this article is available online.

References

Afesorgbor, S. K., & Demena, B. A. (2022). Trade openness and environmental emissions: Evidence from a meta-analysis. *Environmental and Resource Economics*, *81*(2), 287–321.

Afesorgbor, S. K., Fiankor, D. D., & Demena, B. A. (2024). Do regional trade agreements affect agri-food trade? Evidence from a meta-analysis. *Applied Economic Perspectives and Policy*, *46*(2), 737–759.

Anwar, A., & Afesorgbor, S. K. (2021). The effect of remittances on financial development: Evidence from a meta-analysis. *SSRN Electronic Journal*, 1–50. https://doi.org/10.2139/SSRN.3971871

Balcells, L., & Kalyvas, S. N. (2014). Does warfare matter? Severity, duration, and outcomes of civil wars. *Journal of Conflict Resolution*, *58*(8), 1390–1418.

Bara, C., & Hultman, L. (2020). Just different hats? Comparing UN and non-UN peacekeeping. *International Peacekeeping*, *27*(3), 341–368.

Beardsley, K., Cunningham, D. E., & White, P. B. (2019). Mediation, peacekeeping, and the severity of civil war. *Journal of Conflict Resolution*, *63*(7), 1682–1709.

Carnegie, A., & Mikulaschek, C. (2020). The promise of peacekeeping: protecting civilians in civil wars. *International Organization*, *74*(4), 810–832.

Costa-Font, J., McGuire, A., & Stanley, T. (2013). Publication selection in health policy research: The winner's curse hypothesis. Health Policy, *109*(1), 78–87.

Demena, B. A. (2015). Publication bias in FDI spillovers in developing countries: A meta-regression analysis. *Applied Economics Letters*, *22*(14), 1170–1174.

Demena, B. A. (2017). *Essays on intra-industry spillovers from FDI in developing countries: A firm-level analysis with a focus on sub-Saharan Africa* [PhD dissertation, Erasmus University, The Hague].

Demena, B. A. (2021). Effectiveness of export promotion programs, ISS Working Paper No. 688.

Demena, B.A. (2024a). Does export promotion enhance firm-level intensive margin of exports? Evidence from a meta-regression analysis. *Journal of Asian Business and Economic Studies, 3*1(4), 250–262.

Demena, B. A. (2024b). Publication bias in export promotion impact on export market entry: Evidence from a meta-regression analysis. *Applied Economics Letters*, (forthcoming), 1–6. https://doi.org/10.1080/13504851.2024.2306185.

Demena, B. A., & Afesorgbor, S. K. (2020). The effect of FDI on environmental emissions: Evidence from a meta-analysis. *Energy Policy*, *138*, 111192.

Demena, B. A., Floridi, A., & Wagner, N. (2022). *The short-term impact of COVID-19 on labour market outcomes: Comparative Systematic Evidence*. Springer.

Demena, B. A., & van Bergeijk, P. A. G. (2017). A meta-analysis of FDI and productivity spillovers in developing countries. *Journal of Economic Surveys*, *31*(2), 546–571.

Demena, B. A., Reta, A., Jativa, G. B., Kimararungu, P., & van Bergeijk, P. A. G. (2021). Publication bias of economic sanction research: A meta-analysis of the impact of trade-linkage, duration and prior relations on sanction success. In P. van Bergeijk (Ed.), *Research Handbook on Economic Sanctions*. (pp. 125–150). Edward Elgar.

de Rouen, K. R., & Sobek, D. (2004). The dynamics of civil war duration and outcome. *Journal of Peace Research*, *41*(3), 303–320.

Di Salvatore, J. (2016). Inherently vulnerable? Ethnic geography and the intensity of violence in the Bosnian civil war. *Political Geography*, *51*, 1–14.

Di Salvatore, J. (2019). Peacekeepers against criminal violence—Unintended effects of peacekeeping operations? *American Journal of Political Science*, *63*(4), 840–858.

Di Salvatore, J. (2020). Obstacle to peace? Ethnic geography and effectiveness of peacekeeping. *British Journal of Political Science*, *50*(3), 1089–1109.

Dixon, W. J. (1996). Third-party techniques for preventing conflict escalation and promoting peaceful settlement. *International Organization*, *50*(4), 653–681.

Doucouliagos, C., & Stanley, T. D. (2013). Are all economic facts greatly exaggerated? Theory competition and selectivity. *Journal of Economic Surveys*, *27*(2), 316–339.

Doucouliagos, C. (2011). How large is large? Preliminary and relative guidelines for interpreting partial correlations in economics (No. 2011_5). Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.

Doyle, M. W., & Sambanis, N. (2000). International peacebuilding: A theoretical and quantitative analysis. *The American Political Science Review*, *94*(4), 779–801.

Escribà-Folch, A. (2010). Economic sanctions and the duration of civil conflicts. *Journal of Peace Research*, *47*(2), 129–141.

Floridi, A., Demena, B. A., & Wagner, N. (2020). Shedding light on the shadows of informality: A meta-analysis of formalization interventions targeted at informal firms. *Labour Economics*, *67*, 101925.

Floridi, A., Demena, B. A., & Wagner, N. (2021). The bright side of formalization policies! Meta-analysis of the benefits of policy-induced versus self-induced formalization. *Applied Economics Letters*, *28*(20), 1807–1812.

Floridi, A., Demena, B. A., & Wagner, N. (2022). A game worth the candle? Meta-analysis of the effects of formalization on firm performance. *Journal of Developmental Entrepreneurship*, *27*(04), 2250026.

Goldfarb, R. S. (1995). The economist-as-audience needs a methodology of plausible inference. *Journal of Economic Methodology*, *2*(2), 201–222.

Gujarati, D. N., & Porter, D. C. (Eds.). (2009). *Basic econometrics* (5th ed.). McGraw-Hill Irwin.

Haass, F., & Ansorg, N. (2018). Better peacekeepers, better protection? Troop quality of United Nations peace operations and violence against civilians. *Journal of Peace Research*, *55*(6), 742–758.

Hadi, A. S. (1994). A modification of a method for the detection of outliers in multivariate samples. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *56*(2), 393–396.

Havránek, T., Stanley, T. D., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W. R., Rost, K., & van Aert, R. C. M. (2020). Reporting guidelines for meta-analysis in economics. *Journal of Economic Surveys*, *34*(3), 469–475.

Hultman, L., Kathman, J., & Shannon, M. (2014). Beyond keeping peace: United Nations effectiveness in the midst of fighting. The *American Political Science Review*, *108*(4), 737–753.

Kathman, J. D., & Wood, R. M. (2011). Managing threat, cost, and incentive to kill: The short-and long-term effects of intervention in mass killings. *Journal of Conflict Resolution*, *55*(5), 735–760.

Kathman, J. D., & Wood, R. M. (2014). Stopping the killing during the "peace": Peacekeeping and the severity of post-conflict civilian victimization. *Foreign Policy Analysis*, *12*(2), 149–169.

Mebratie, A. D., & van Bergeijk, P. A. G. (2013). Firm heterogeneity and development: A meta-analysis of FDI productivity spillovers. *The Journal of International Trade & Economic Development*, *22*(1), 53–74.

Murshed, S. M. (2009). *Explaining civil war*: A *rational choice approach*. Edward Elgar Publishing.

Phayal, A. (2019). UN troop deployment and preventing violence against civilians in Darfur. *International Interactions*, *45*(5), 757–780.

Phayal, A., & Prins, B. C. (2020). Deploying to protect: The effect of military peacekeeping deployments on violence against civilians. *International Peacekeeping*, *27*(2), 311–336.

Regan, P. M., & Aydin, A. (2006). Diplomacy and other forms of intervention in civil wars. *Journal of Conflict Resolution*, *50*(5), 736–756.

Regan, P. M., & Meachum, M. S. (2014). Data on interventions during periods of political instability. *Journal of Peace Research*, *51*(1), 127–135.

Sambanis, N., Skaperdas, S., & Wohlforth, W. (2020). External intervention, identity, and civil war. *Comparative Political Studies*, *53*(14), 2155–2182.

Sawyer, K., Cunningham, K. G., & Reed, W. (2017). The role of external support in civil war termination. *Journal of Conflict Resolution*, *61*(6), 1174–1202.

Sousa, R. R. P. (2014). *Effect of external interventions on conflict intensity* [doctoral dissertation, PhD-értekezés, Rotterdami Erasmus Egyetem].

Stanley, T. D. (2005). Beyond publication bias. *Journal of* Eco*nomic* S*urveys*, *19*(3), 309–345.

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, *70*(1), 103–127.

Stanley, T. D., & Doucouliagos, C. (2012). *Meta-regression analysis in economics and business*. Routledge.

Tukey, J. W. (1977). Exploratory Data Analysis. Addison—Wesley.

van Bergeijk, P. A. G., Demena, B. A., Reta, A., Jativa, G. B., & Kimararungu, P. (2019). Could the literature on the economic determinants of sanctions be biased? *Peace Economics Peace Science and Public Policy, 25*(4), 20190048. https://doi.org/10.1515/peps-2019-0048

Wood, R. M., Kathman, J. D., & Gent, S. E. (2012). Armed intervention and civilian victimization in intrastate conflicts. *Journal of Peace Research*, *49*(5), 647–660.