reason I will argue that data's validity is maximized by using the most detailed data available.

Duration data are from Gleditsch *et al.* (forthcoming), based on Gates & Strand (2006). Start and end dates are here coded as precisely as possible.[23] Data differentiates between 'major' and 'minor' start and end dates. The former refers to the first and last date of violence, while the latter is the first and last dates where the UCDP/PRIO Armed Conflict definition of conflict is fulfilled. In line with my above definition of internal conflict, I use the latter.

Conflicts in the data that were initiated before 1946 or was still going on at the end of 2004 are censored. In my data, seven conflicts had broken out prior to January 1, 1946, and 28 conflicts were still running at the dataset's end date of December 31, 2004.

4.3.3 Measuring Religious Cleavages

The main explanatory variable concerns religious cleavages. This deserves extra attention. When operationalizing religious cleavages, there are two main alternatives. First, the issues of the conflict can be analysed, searching through documents and statements for the rebels' goals. In this case, a religious cleavage can be defined as a central religious incompatibility in the conflict. According to this definition it would be a religious conflict when the rebels actively seek a religious change of some sort, either by establishing or disestablishing a state religion, fighting religious discrimination, seceding to form a new state based on religious principles other than those in the existing state, or simply removing the 'infidel' incumbents. Second, the focus can be on the identity of the involved actors. This is the more common of the two alternatives (Nordås 2007: 8). In this case, there is a religious cleavage when the two parties belong to separate religions, when they adhere to different denominations of one religion, or when one side is religious and the other side not. Obviously the best operationalization is to include both issue- and identity-based cleavages. However, considering the lack of comprehensive sources and the temporal limitations of this

---

[23] The dataset also provides precision scores for each start and end date. For more on this, consult Gates & Strand's (2006) dataset and its readme-file, both available from http://www.prio.no/CSCW/Datasets/Armed-Conflict/Onset-and-Duration-of-Intrastate-Conflict/Duration-Data-v1-2006b/.

thesis, collecting data on both issues and identities is too great a task. Therefore, a choice must be made. As indicated in the introduction, the choice has fallen on the identity-based definition. This choice needs justification.

First, are the two alternatives equally valid operationalizations of the concept religious cleavage? The former definition is the only one that catches whether or not religion actually is a cause or main element of the conflict. A struggle between parties of different religions might have nothing to with religion *per se*. In fact it may be exclusively about politics or economics. Thus, the former might be considered a more valid operationalization of the underlying concept (*ibid.*: 8). Nevertheless, an identity-based definition may be better able to capture several important aspects. For instance, it is easier to establish a division of 'us' and 'them' when there is an identity-based cleavage. This is both because the groups have different belief-systems and because the division is more salient. As a consequence demonization and dehumanization is more likely to occur. Furthermore, the notion of religious institutions as arenas for mobilization may be more important in identity-based religious conflicts. In comparison, a religious leader trying to mobilize his community will more easily meet internal opposition among followers where there is an issue-based cleavage and not an identity-based one. Concerning duration, intergroup trust will arguably be easier to establish if there is only an issue-based cleavage as religious authorities on both sides may be respected by supporters of both sides.

The two approaches do occasionally diverge on the question of religious cleavage. An example is Algeria, where radical Sunni insurgents fight moderate, secular Sunni incumbents. An issue in the conflict is the role of religion in public life (Juergensmeyer 1993: 48, 168).[24] Still, the parties do not diverge on religious adherence; they are both Sunni Muslims. The identity-based approach here fails to capture the cleavage.

Second, the two approaches can be said to differ with respect to reliability. In order to assess whether or not there is an issue-based religious cleavage, it is necessary to search through a vast amount of speeches and scholarly work. This is a huge task

---

[24] It should be noted, however, that the role of religion in the Algerian conflict is contested (see for instance Schulhofer-Wohl 2007). The disagreement over the relevance of religious issues in Algeria demonstrates the inherent difficulties of an issue-based definition.

and different researchers do not necessarily have access to the same material. Additionally, it might also be necessary to interview central actors where text sources are scarce. A problem will occur where some recognize the conflict as a religious one while others claim that faith was of no importance. The ensuing coding process is to a considerable extent contingent upon the discretion of the coder. As a further complication, the use of religious rhetoric in arguments and appeals does not necessarily imply that religion is an important issue. This is where the identity-oriented approach has one of its strengths. Although not always crystal clear, the religious composition of the parties is more readily available, and whereas coders might assign different weights to different issues, they will arguably code religious affiliations more consistently. Additionally, information on identity is less likely to be biased. Many groups have an explicit religious profile. Those who do not are sometimes based in a certain region where most people adhere to a certain faith. This makes coding easier and more reliable for the latter approach.

A third and related issue is that data on identities are easier available than issue-data. This is because the identities of belligerents are more widely known than their goals. In order to know the objectives of a rebel group, you first need to know the group itself. Moreover, for many conflicts, especially those farthest back in time and the short-lived ones, little documentation is available and easily accessible. This has consequences for reliability.

Fourth, it is a question of costs. Examining the issues in a conflict may be highly time-intensive. And if interviews are needed because of scarce written sources, it may involve costly and time-consuming travels. Furthermore, it may be risky if one is to travel to conflict zones. In sum, data's validity for the identity-based definition is considered to be at least as high as for the issue-oriented definition. Moreover, the former suffers less from poor availability and high costs.

In order to identify religious cleavages, I have first coded the religious affiliations of the relevant groups and then decided whether or not fighting parties differ from each other. Conflicts with such a difference have been assigned the value 1, indicating the presence of a religious cleavage. Cases without a religious difference have been given the value 0. I have been open for a variety of religious traditions. In

total, 15 distinct values have been used, including a value for non-religious groups.[25] The assignment of values has been based on the following criteria. First, the government side was assigned a value based on the religious identity of the government or governing party as a whole when this was known. For instance, Israeli governments were coded Jewish. Where this was unclear it was examined whether the regime had a distinct ethnic or regional identity. If this was the case and the ethnic or regional group had a distinct religious affiliation, the regime was assigned a value according to this. Where the coding still was undetermined a value was assigned based on the head of government's religious adherence.

Similarly, rebel groups have been assigned a value based on their adherence where this was obvious. For instance, the Somali al-Itihad al-Islami (AIAI) insurgents were coded Sunni and the so-called Sikhist insurgents in India were coded Sikhist. The remaining units were given values based on ethnicity or region where this was distinct and related to a specific religious affiliation – for example the Tibet insurgency in China has been coded 'Other/Mixed Buddhist' based on the beliefs of the vast majority of Tibetans (Tibetan Buddhism) – or based on the group's leaders. The latter is the case with, for example, coups d'etat, where the identity of followers usually is unknown. An example is the attempted coup d'etat in Cameroon in 1984. The rebels were supporters of former president Ahmadou Ahidjo. Ahidjo was a Sunni Muslim; hence the rebels are treated as Sunnis.

In several cases the rebel side was made up by more than one group and these did not always adhere to the same faith. In order to code a joint value to the rebel side, the religion of the plurality of the groups was assigned, disregarding the relative size of the groups. It would be favourable to take relative size into consideration, but such information is unreliable and unavailable for a number of groups. Where two religions were equally frequent, the UCDP/PRIO Armed Conflict Dataset (Gleditsch *et al*. 2002) and the COW Intra-State Wars dataset (Sarkees 2000) were used to identify the main group. The rebel side has then been assigned the value related to this group.

---

[25] These are: Protestant, Catholic, Orthodox, Other/Mixed Christian, Sunni, Shi'a, Other/Mixed Muslim, Jewish, Hindu, Sikhist, Theravada, Other/Mixed Buddhist, Animist, Non-religious, and Other. In addition, I have considered values such as Jainist, Zoroastrian, Mayahana Buddhist, Shinto, Taoist, and Confucian.

In some cases the government is replaced in the course of the conflict or a rebel group joins or leaves an ongoing struggle, thus influencing the presence of religious cleavages as defined.[26] For the duration analysis, each conflict has been assigned one value for the entire conflict, based on the main actors. For the intensity analysis the value of this variable is allowed to change from year to year, something that is needed as different conflict-dyads are not differentiated (see section 4.2.1). Where a government fights several separate rebellions, the cleavage variable is assigned the value that is most frequent among these rebellions. Where the values (0s and 1s) are equally frequent, the unit is treated as one with a religious cleavage.

Not all people are religious believers. In the same manner as a cleavage may exist between different religious groups, a religious group and a non-religious group might see each other as adversaries. The same effects could, therefore, be expected where atheists fight a group with a distinct religious adherence. Communism is ideologically opposed to religion and communist regimes have sought to eradicate organized religion (Barro & McCleary 2005: 1344). According to Bukharin & Preobrazhensky, '[s]cientific communism, in its judgements concerning natural phenomena, is guided by the data of the natural sciences, which are in irreconcilable conflict with all religious imaginings' (1969: 300). Consequently, '[r]eligion and communism are incompatible, both theoretically and practically' (*ibid*.: 299) and a Communist that 'continues to cling to his religious faith … ceases thereby to be a communist' (*ibid*.: 300). Ergo, Communist governments and rebel groups are coded non-religious.

A variety of sources – publications, online sources and personal communication – have been consulted. Among the most frequently used sources are publications from DeRouen & Heo (2007) and U.S. Department of State (2008); datasets from Fox (2008), Roeder (2003), and Svensson (2007a); as well as HighBeam Encyclopedia (2008). These are all considered highly reliable.[27] Where these proved insufficient, I

---

[26] In order to assess whether or not this influenced the findings, an alternative coding based allowing for varying values has been included. The results did not change significantly. These results are not reported.

[27] HighBeam Encyclopedia (2008) is a collection of articles from reliable encyclopedias such as Britannica and Columbia as well as various newspapers and magazines. As Svensson (2007a) does not include a non-religious category, information from his dataset is not used separately without a wide search for such information. The utilized datasets are available at the following URLs: Fox (2008) at

have consulted other publications or contacted country experts. Additionally, some online sources considered less reliable have been consulted. The latter are not regarded sufficient alone; rather they are used for verification purposes. For a detailed listing of sources used in different cases, see appendix A and the associated reference list.[28]

Missing data is generally a problem for studies of violent conflict. This is so also for this project. I have been unable to identify whether or not a religious cleavage was present in 34 conflicts in 18 different countries. This leaves me with 241 conflicts for the duration analysis and 1,035 conflict-years for the intensity analysis.

4.3.4 Religious Discrimination

This variable is intended to capture state discrimination against some or all religions. Data on religious discrimination is taken from Jonathan Fox's Religion and State (RAS) dataset.[29] Two RAS indicators are utilized here. First, there is an indicator measuring the degree of discrimination against minority religions on a scale going from 0 to 48. This is a composite variable summing up the values for 16 different types of discrimination, each measured on a scale going from 0 to 3 (Fox 2004a: 5-7). The second indicator measures the degree of discrimination against the majority or all religions. This is included to get a grasp of religious discrimination in general, such as may be present in Communist countries, and is 'qualitatively different from restrictions on minority religions' (*ibid*.: 7). The second indicator goes from 0 to 33, and is a composite of 11 types of discrimination measured on a four-level scale similar to that of the first indicator (*ibid*.: 7-8). Values on the two indicators have been added into a single additive index, giving a total range of 0 to 81 where 81 is maximum discrimination and 0 represents no religious discrimination. This seems more reasonable than to weigh the two, as they both are made from a sum of variables measured on a similar scale.[30]

---

http://www.biu.ac.il/soc/po/ras/downloads.html; Roeder (2003) at http://weber.ucsd.edu/~proeder/data.htm; and Svensson (2007a) at http://jcr.sagepub.com/cgi/content/full/51/6/930/DC1.

[28] Throughout the coding process I have made an effort at triangulation of sources. In a few cases where only one source has been found, coding is based on this if the source is considered highly reliable and no contradictory information is found. Where information from the highly reliable sources is missing, unclear, or contradictory, values are treated as missing.

[29] Available at: http://www.biu.ac.il/soc/po/ras/downloads.html.

[30] For a detailed account of the included types of discrimination, see Fox (2004a).

This is considered a highly valid indicator of religious discrimination. It includes a broad range of discriminating acts and policies. Reliability is also considered to be high. Throughout the process of coding, the project director has supervised coders to ensure that they use the same methodology and criteria. Furthermore, about one fourth of the states have been recoded by another coder to ensure objective coding (*ibid*.: 1).

Fox's data cover the period 1990-2002. Since changes from year to year are relatively rare, and where they occur they tend to be marginal, I have extrapolated the data through the entire period covered by my data. Four countries, constituting 23 conflict-years, still have missing values after this operation, as these ceased to exist prior to 1990.[31] These have been assigned the variable's mean value.[32] These operations increase the prospects for significant results. Still, such operations may have unfortunate effects. In order to assess the chances for skewed results, the stability of discrimination values have been examined. Results show that a vast majority of countries have unchanged values throughout the twelve-year period (see appendix B).

This indicator has been centred to minimize violations of normality. As religious discrimination might impact differently on religious and non-religious conflicts an interaction term is also included.

4.3.5 Religious Legitimacy

Religious legitimacy is hard to measure directly. A proxy should account for whether or not religion is used in public debate. Following Fox (2000a) I use the presence of an official state religion as an operational definition. Values for each year has been decided based on whether or not a state religion was present at the beginning of the year. The variable is dichotomous. Units with a state religion are given the value 1, while those without a state religion are assigned the value 0. Data is collected from

---

[31] These are Arab Republic of Yemen, People's Republic of Yemen, Muscat and Oman, and Republic of Vietnam (South Vietnam). In sum, these make out six conflicts and 24 conflict-years.

[32] As mentioned below, the same has been done for some of the control variables. For each variable where this is done, a dummy variable has been added to control for potential effects from this operation. For the dummy variable, units with an original missing value are coded 1 and other units are coded 0. For more on this, see section 4.4.

Barrett (1982), Barrett *et al.* (2001), Nordås (2004b), and U.S. Department of State (2008).

The validity of this variable requires discussion. Religious legitimacy can be said to consist of more elements than whether or not there is an official religion. Still, this does to a large degree capture how legitimate religion is in the polity as a whole. Moreover, it captures the fact that Communist states, which are ideologically opposed to religion, are atheist. The main problem lies at the local level. Even in Communist states there may be smaller communities where religion is considered highly legitimate in public debate, and in such communities religion might be used in the mobilization process. An advantage of this indicator is its reliability. The presence of state religions is normally obvious and not dependent on individual discretion. More detailed proxies, considering local communities and other aspects of religion in public life, would arguably suffer under reliability. In sum, data's validity is considered acceptable. This variable has no missing values.

An interaction term is included as religious legitimacy may affect religious and non-religious conflict differently.

4.3.6 Religious Demography

In order to capture different aspects of religious demography one index of religious fractionalization and one of religious polarization is included. Formulas for the indices are taken from Montalvo & Reynal-Querol (2005). The former is based on the following formula:

$$FRAC = 1 - \sum_{i=1}^{N} \pi_i^2 = \sum_{i=1}^{N} \pi_i(1 - \pi_i)$$

where π_i is the proportion of people that belong to the religious group *i* and *N* is the total number of groups (*ibid.*: 797). This index is interpreted as the probability that two randomly drawn people from a given population belong to different religious groups.

Religious polarization is calculated from the following formula:

$$RQ = 1 - \sum_{i=1}^{N} \left(\frac{1/2 - \pi_i}{1/2} \right)^2 \pi_i = 4 \sum_{i=1}^{N} \pi_i^2 (1 - \pi_i).$$

This index is meant to capture how far a population is from a bipolar distribution (*ibid.*: 798). Both indices range from 0 to 1, where 1 is maximum fractionalization and maximum polarization, respectively. Montalvo & Reynal-Querol (*ibid.*) also include scores on religious fractionalization and polarization in most of the countries in my data. However, their classification of religions differs from the one used to measure religious cleavages in this thesis. For instance, their data do not differentiate between different branches of Christianity. Hence, the scores are not entirely commensurable. For this reason I choose to use data from the RAS dataset, which measures the proportion of populations adhering to 16 different religious traditions. These denominations are very similar to those included in my classification scheme.[33] The RAS data is then used to calculate scores of religious fractionalization and polarization, using the above formulas.

After adding data that suits my needs, both indices are considered valid operationalizations. The RAS data are collected from Barrett (1982), Barrett *et al*. (2001) and the CIA World Factbook (Fox 2004c), sources generally considered highly reliable and widely used. Only Yemen (including the former North and South) has missing values. This has been corrected using information on the country's demographics from U.S. Department of State (2008) and NationMaster.com (2008).

As for legitimacy a problem is that religious fractionalization and polarization are macro measures relating to the population as a whole, not only to the fighting parties or their local communities. When belligerents recruit soldiers they may focus on a particular region, and the rest of the country is then of little relevance. This points to the importance of geography and scope in civil wars. For the time being, more

---

[33] The RAS dataset includes Catholics, Protestants, Orthodox Christians, other Christians, Sunnis, Shi'ites, other Muslims, Buddhists, Hindus, Jews, Animists, Confucians, Sikhs, Bahá'i, non-religious, and others for 174 states and areas. Additionally, different types of Protestantism are separated. In my data I have used aggregate numbers of Protestantism as this makes the classification comparable to the coding of cleavages.

specified indicators are not available.[34] As characteristics of the population are expected to influence conflict dynamics, these indicators are used as the best available proxies.

These indices have been centred to minimize violations of normality. An interaction term is included for religious cleavages and each of the two variables on religious demography. A squared term for polarization is included in the duration analysis.[35]

4.3.7 Control Variables

Control variables are included in order to hold possible confounding factors constant and ensure that findings are not simply due to spurious relationships. With the introduction of control variables the impact of the explanatory variables can be evaluated for specific control groups (Skog 2004: 44). This breaks the bivariate correlations down into partial correlations, and only then can we establish actual impacts.

First, an indicator of regime type is included. Democracies have legitimate and institutionalized manners to resolve conflict and are better able to respond to challenges without resorting to massive use of violence. This should make conflicts in democracies less intense than in other countries, because 'autocracies and non-democratic new nations, on the other hand, typically follow policies of deadly response to both protest and rebellion' (Lichbach & Gurr cited in Benson & Kugler 1998: 198, n. 1). Furthermore, democratic institutions should theoretically be better able to end civil war through negotiations and non-violent means. In comparison, autocratic regimes lack credible institutions that might induce trust between the parties. Data on regime type are taken from Gleditsch *et al*. (forthcoming). The chosen indicator is based on the Polity IV project. The Polity index scores regimes on a 21-point scale, ranging from +10 to -10. The score is found by subtracting the value on an autocracy scale (0-10) from the value on a democracy scale (0-10) (Marshall & Jaggers 2000:

---

[34] A group of researchers at PRIO are working to improve such indicators to consider geography, scope, and local factors. See for instance Buhaug & Gates (2002), Buhaug & Lujala (2005), Buhaug *et al.* (2008), and Rød & Buhaug (2008).

[35] Squared terms have been tested for both indicators in both analyses, but only this proved significant.