Online Appendix for "Segregation, Polarization, and Ethnic Conflict"

July 29, 2014

This online appendix accompanies the paper. In section A.1 we explain how the dynamic model of conflict presented in the paper is simulated. In Section A.2, we provide more details on the implementation of the Schelling model in NetLogo and how we use it to generate hypothetical countries with different levels of segregation and polarization. Section A.3 provides additional details on the illustrative results discussed in Section 4 in the main text. In Section A.4, we present our simulation results for conflict outcomes other than attacks – presented in the main text – namely total effort, the number of victories, and conflict duration. While in the paper we showed that a unique analytical result between segregation, polarization, and conflict outcomes cannot be derived, in Section A.5 we generate an injective map between regional population shares and segregation, and consequently between segregation and first period conflict, when regional population shares are restricted in a particular way. Section A.7 gives the summary statistics for the Indian and Bosnian data we use in the empirical analysis. Finally, Section A.8 presents and discusses robustness checks of our main empirical results.

A.1 Simulating conflict

We perform simulations of the model presented in Section 2 of the main text in the following way. Figure A1a gives an example of a country with 25 regions, 50 percent of each ethnic group (maximum polarization), and segregation level of 0.41 in this country (i.e. the groups are quite mixed throughout the country). For each region, the population share of the dissatisfied group is indicated by a number and by a corresponding shade of gray, where darker shades represent more of the dissatisfied group.

In the first period, members of the dissatisfied ethnic group in each region decide whether to attack the other group. As explained in Section 2, they will attack whenever they represent more than 61.8 percent of the population. These attacks are indicated with a "1" in Figure A1b. When the dissatisfied group attacks, they may win or lose, as determined by their equilibrium probability of winning.[1] When the region is won, it turns completely black and the number of attacks indicated for that region in future periods will remain the same because we assume that

[1] As mentioned earlier, the realization of the equilibrium probability of winning is determined based on a draw from a uniform distribution on [0,1] and whether this draw is less or greater than the equilibrium probability of winning. For example, if the equilibrium probability of winning is 80% and the random draw is anything less than 0.8 a victory is assigned. The probability that a random draw from a uniform distribution on [0,1] is less than 0.8 is exactly 80%. Hence, the method we use is just a statistically correct way to go from equilibrium probabilities to victories and losses consistent with these equilibrium probabilities.

victories of the dissatisfied ethnic group are final. If the dissatisfied group loses, it can re-attack in the following period as long as the condition for attack is satisfied.[2]

At the end of the first period, s_{1j} is calculated for each region j as the number of dissatisfied group's victories divided by the total number of attacks in the regions neighboring region j. For example, the region in the far bottom right corner has 56 percent of the dissatisfied group which is not enough to attack in the first period. But in the second period, they will attack because their $s_{1j} = 1$, so they only need $n_{1j} > 0.448$ to attack, which is satisfied.[3] This demonstrates how, due to neighborhood effects, conflict may spread to regions where it would not have occurred otherwise. In each period, s_{1j} and p_{1j} are recalculated to determine the next period attack in each region. The simulated conflict game will continue as long as the dissatisfied group in at least one region has a probability of winning greater than 0.618, and this region has not yet been captured.[4,5]

Due to the probabilistic nature of winning, Figures A1b-A1e represent just one possible way in which the conflict could evolve in this hypothetical country. The conflict may be associated with different amounts of expended effort, different number of attacks, captured regions, and could end in fewer or more periods, depending on the exact realization of the equilibrium probability of winning in each region where the dissatisfied group attacks. This is another reason why we resort to simulations. However, we show in Figures A8, A9, A10, and A11 that this aspect of the model does not in any way change the interpretation of our comparative statics. Rather, the probabilistic nature of the dynamic model only shifts the mean of the outcomes at any given combination of segregation and polarization, but not the overall relationship. This is as expected, since the path-dependency is due to a random draw which we use to operationalize the probability of winning, rather than any systematic component of the model. In other words, the path-dependency arises

[2]Note that the fact that the dissatisfied group can re-attack does not necessarily entail an eventual victory for the dissatisfied group – they may lose in one period and find that the condition to attack is no longer satisfied in the future periods. On the other hand, the permanency of victory is a simplification, but one that allows to more clearly focus on the spread of conflict, which is always finite in our simulations. The post-conflict arrangements are beyond the scope of this study.

[3]This happens because their only neighbor who attacked in the first period, the diagonal neighbor with 90 percent of the dissatisfied group, won the battle, so the share of co-ethnic neighbors' successes is 1.

[4]Our model is therefore static in the sense that in each region, ethnic groups maximize their utility myopically, i.e. only in the present period. However, the model is dynamic in the sense that there is a state variable that evolves over time and depends on past actions of neighbors. Alternatively, we could set this problem as a dynamic game with forward-looking agents who take expectations over the states tomorrow when deciding whether to attack and what effort level to choose. However, with forward looking agents, a region would have to take into account N state variables, one for each of N regions. This is because a region is affected by its own neighbors, but to take into account the future actions of one's neighbors, one needs to take into account the future actions of their neighbors, etc., until the full set of everybody's neighbors is exhausted. As the number of state variables increases, the model becomes prohibitively difficult to solve even computationally for anything but a very low N. On the other hand, reasonable simplifications would take us too far afield from our focus here.

[5]The reader might wonder if in a model with rational forward-looking agents it might happen that there is no conflict because of free-riding. We argue that this is unlikely. Suppose an agent representing a region has an incentive to attack today, but decides to wait because it would be even better to attack after his neighbors have won. This forward-looking leader would also know that his neighbors are thinking the same and that this would lead to no attacks anywhere. But then going back to step 1, the forward-looking leader would realize that his choice is not between attacking now or attacking later, but between attacking now or never attacking. Then it is better for the leader to attack in the first period, and hence, no rebellions anywhere is not an equilibrium. This argument is exactly correct if the segregation level in all regions is identical, so that their decisions to attack are symmetric. However, even if the regions are not all identical in terms of segregation, if the condition for attack is satisfied in the first period, it cannot be an equilibrium to wait for others to attack and possibly end up with no region attacking – rational forward looking agents would anticipate this and hence attack in the first period.

because there is often some uncertainty as to whether the dissatisfied group will win following an attack (i.e. the probability of winning is often lower than one).

A.2 Schelling Model to Generate Segregation

The Schelling model has one key parameter common to all agents – the percent of their immediate neighbors they desire to be of the same racial (or ethnic) group as themselves, or "percent similar wanted." This can be viewed as a segregation parameter.[6] If an agent's immediate neighborhood contains less than their percent similar wanted, that agent can move in any direction until it finds the satisfactory neighborhood. If all agents find a suitable location, a steady state is reached. This steady state is characterized by a certain level of spatial segregation.

On average, when the percent similar wanted is high, the two groups will tend to be more segregated than when the percent similar wanted is low. To generate different segregation levels, we vary the percent similar wanted between zero and 87.5 percent in increments of 4.17 percent.[7] Note that different runs of the same Schelling model generate different geographic distributions, even with the same percent similar wanted. This is because the agents can move in different order and in different directions in a two-dimensional world and are initially scattered in random patterns.

After the steady state has been reached in each run of the Schelling model, we split the Schelling world into square regions (five by five slots), which can contain at most 25 agents, or a maximum of one agent per slot.[8] The Schelling world that we use is 55 by 55, i.e. has 3,025 slots for agents to occupy. There are between 1,300 and 1,700 agents, depending on the polarization level.[9] When there are only two groups, the polarization level can be calculated directly from the relative size of the groups.

In the paper, we focus on eight levels of polarization: 1, 0.97, 0.93, 0.85, 0.75, 0.55, 0.30, and 0.20. These polarization levels correspond to the following countrywide population shares (0.5,0.5), (0.42,0.58), (0.37,0.63), (0.31,0.69), (0.25,0.75), (0.17,0.83), (0.08,0.92), (0.05,0.95). These levels are chosen because they nicely demonstrate the changing role of the interaction between polarization and segregation in conflict outcomes.[10] In each region, we calculate the population share of the

[6] The model was originally motivated by racial residential segregation in the United States, and the hypothesized theory that after the share of African-Americans in a neighborhood reaches a certain threshold, called the tipping point, the white residents move en masse to another area which has a higher share of white residents.

[7] This increment is equal to $\frac{1}{24}$. The default size and shape of NetLogo neighborhood is a square containing nine single-agent slots, or patches, with the calling agent in the center. We enlarge this square neighborhood to include 25 patches. This means that any agent has 24 neighbors, and so, there are essentially only 24 relevant values of the percent similar wanted. 87.5 percent, or $\frac{21}{24}$ is the highest percent similar wanted that allows the steady state for a given polarization level in our implementation.

[8] In the Schelling model, some slots must be vacant so that agents can move to a new location if they find their current location undesirable. This means that some of the regions we define will have all the slots occupied (i.e. 25 agents), and others may have empty slots.

[9] Our aim is to produce as similar a range of segregation levels for each polarization level for the sake of comparison. For a fixed size of the Schelling world, whether the steady state of the model is reached depends on the number of agents, proportion of the two groups – polarization, and the percent similar wanted. In general, a higher percent similar wanted is needed to achieve the same level of segregation when the two groups are unequal (low polarization) compared to two groups of similar size (high polarization). But for a given number of agents, simulations with a very high percent similar wanted often fail to reach the steady state, as the agents are never fully satisfied and the model runs indefinitely. Reducing the number of agents avoids this problem.

[10] It is possible to produce near-continuous variation in polarization levels; however, this is immensely computationally intensive. Moreover, as discussed in the previous footnote, setting the model parameters to get a full range

ethnic groups by dividing the number of agents in that ethnic group by the total number of agents in the region.[11] We then use this output as the input for the conflict simulation game, which is run on an 11 by 11 world comprising 121 regions.

A.3 Illustration: Further Results

Here, we provide further details on the illustrative results discussed in Section 4. We run 1,000 repetitions of the dynamic model of conflict for the hypothetical countries in Figure 1. Figure A2 presents the distribution of attacks and victories for these simulations – attacks are shown in black and victories in gray.[12] Table A1 gives the summary statistics for all the outcomes.

In Table A1 we see that in the high-polarization and low-segregation country the median number of regions won is 69 (even though there are only 20 attacks in the the first period). In contrast, in the high-segregation country, the median number of captured regions is 62 (even though there are many more attacks in the first period, 59). Hence, when polarization is high, despite more contained initial breakout of conflict in the low-segregation country, there are eventually more attacks and victories. Overall, when polarization is high, the highly segregated country has smaller variance in attacks and victories – a battle either does not occur, or it is quickly won by the attackers (because most regions have either a clear majority or a clear minority of the dissatisfied group). In the case of the low-segregation country, the average number of attacks and victories is higher and the variance is higher, because of the dependence on neighbors' success and location of previous attacks. When polarization is low, increasing segregation has the opposite effect – there are more attacks in the more segregated country. The effect of segregation, therefore, is *conditional* on the level of polarization, even though the underlying dynamic of conflict is identical.

Figure A3 presents the distribution of total effort expanded in conflict for the 1,000 simulations. The results are quite similar to the results for total attack and victories. In the highly polarized country, total effort decreases with segregation. In contrast, in the low polarization country, total effort increases with segregation.[13]

The logic for effort is the same as for attacks and victories. In the highly polarized country, effort is high when segregation is low because there are many regions to which conflict spreads, where it is difficult to win because of low population shares (around 50 percent) of the dissatisfied group. On the other hand, when segregation is high regional battles will not require a lot of effort because members of the dissatisfied group are fighting only against a small local minority of satisfied group (or not fighting at all when they themselves are in the minority). The story is reversed in the low-polarization country, since low segregation now means that the share of the dissatisfied group in each region is only about 5 percent (close to the countrywide share of the dissatisfied group), hence the conflict never begins. But in the example when segregation is high, there are 27 regions with 100 percent dissatisfied group, 1 region with 66 percent and 2 regions with 50 percent. Now,

of segregation levels for each polarization level is non-trivial.

[11]In a rare case that the entire region is empty, the shares equal the proportion of groups for the entire country.

[12]Obviously the number of attacks is greater than or equal to the number of victories, because some attacks are unsuccessful.

[13]Note that in any country with perfect segregation, i.e. in which each region has only members of one group, total effort must be zero. According to the model, if region j consist 100 percent of members of the dissatisfied group, then the probability that they capture region j is $p_j = 1$ and consequently total effort $E_j = \sqrt{\frac{1}{2} p_{1j}(1 - p_{1j})} = 0$. If every region has either zero or 100 percent of the dissatisfied group, then total effort is zero.

the 27 regions are instantly captured by the dissatisfied group with no effort, and some effort is expended in the other 3 regions, once the conflict spreads there.

Figure A4 presents the distribution of conflict duration. It is again clear that for high-polarization countries, higher segregation is associated with longer conflict. When looking at the two low-polarization countries, the pattern is reversed, and conflict is longer in the high-segregation country.

Finally, let us also mention the extreme cases. Consider a country which has maximum polarization and maximum segregation. This would mean that the country is split in two parts, each part perfectly homogenous with either the satisfied or the dissatisfied group. In this case, conflict would almost be senseless because there would be no members of the other group to fight against in either part of the country. At the other extreme, suppose there is a country such that the population shares of the two groups are exactly 50 percent in each region as in the country overall. Now, our model suggests that conflict would never begin, because in no region would either group have more than 61.8 percent population share which is needed for attack. These illustrations indicate that segregation can play an important role for the initiation and spread of conflict.

A.4 Extended Simulation Results

In the main text, we focus mainly on one outcome – the number of attacks by the dissatisfied group. Conflict simulations confirm our expected findings – the effect of segregation depends on the level of polarization. When polarization is very high, higher segregation is associated with *less* conflict. Then, as polarization decreases the effect of segregation begins to change, so that at some point increasing segregation neither increases nor decreases conflict. When polarization is at lower levels, increasing segregation is associated with *more* conflict.

At the maximum level of polarization, shown in the top left corner of Figure 2, the number of attacks is decreasing with segregation almost linearly, except at extremely low segregation. At minimum segregation, the number of attacks is zero. This is because regional and countrywide shares are equal at minimum segregation, so the dissatisfied group represents 50 percent in every region and is thus below the critical threshold for attacks derived in the model. Next, over a very narrow range of segregation levels, from zero to about 0.07, the number of attack soars to 100. As long as conflict begins in some region, it eventually engulfs almost the whole country because the dissatisfied group is a significant presence in almost all the regions.

At somewhat lower polarization levels of 0.97, 0.93, 0.85 (or 42%, 37% and 31% of the dissatisfied ethnic group, respectively), all shown in the first row or Figure 2, the plots change quite dramatically. Now the relationship between segregation and the number of attacks is no longer downward sloping, and over a large range, segregation neither increases nor decreases the number of attacks. The number of attacks at low segregation levels is no longer soaring as in the case of maximum polarization, but is increasing more slowly and over a wider range of segregation – from zero to about 0.2 when polarization is .97, zero to 0.3 when polarization is 0.93, and zero to about 0.4 when polarization is .85. Thereafter, the number of attacks either levels off and remains flat (when polarization is 0.97 and 0.93), or continues to rise with segregation, as is the case when polarization equals 0.85.

At lower polarization levels, shown in the bottom row of Figure 2, the number of attacks is typically zero when segregation is low. At polarization levels of 0.75, 0.55, 0.3, and 0.2 (or 25%, 17%, 8% and 5% of the dissatisfied ethnic group, respectively) the number of attacks increases with segregation – the opposite from the effect of high polarization, and markedly different from the effect at intermediate polarization levels.

Below we also present our simulation results for other conflict outcomes, namely total effort, the number of victories, and conflict duration. These are shown in Figures A6, A5, and A7, respectively. It is interesting to now look at Figure A5, which plots the relationship between segregation and the number of victories, by polarization level. Looking again at maximum polarization (top left), we observe a fairly linear, slightly downward sloping curve, but with a slope that is much less than in the case of attacks. The slope is in fact so flat, that the number of victories at the end of conflict is only slightly lower when segregation is high.

In Figure A6, which plots effort, we also see a linear downward sloping line that reaches zero effort at maximum segregation, as the model predicted.

Moreover, it is instructive to note that there are departures from linearity, especially at polarization levels of 0.97 and 0.93. The outcomes are often hump-shaped, starting from zero at minimum segregation and falling back to zero at maximum segregation. What varies is the location of the peak – it shifts to higher segregation levels as polarization decreases.

Figures A8, A9, A10, and A11 demonstrate that our comparative statics are not affected by the stochastic element of the dynamic model, discussed in Section 2.1. The results are quite similar to those for the median (shown in red, which are equivalent to those shown above) when analyzed at other deciles of the distribution produced by our simulations.

A.5 Special case: Injective function between population shares and segregation

We showed in the main text that a unique analytical result between segregation, polarization, and conflict outcomes cannot be derived, and that therefore we need to resort to simulations in Section 5 to derive comparative statics. However, it is possible to generate an injective map between regional population shares and segregation, and consequently between segregation and first period conflict, if regional population shares are restricted in a particular way so that they all depend on a single variable, q. We will consider two cases: first, a country with polarization equal to 1, i.e. in which each group has 50% of the total country population ($n_1 = n_2 = 0.5$), and second, a country with polarization of 0.19, i.e. in which the dissatisfied group has only 5% of the total country population ($n_1 = .05, n_2 = 0.95$). Suppose each country has $j = 1, ..., J$ regions, and each region has the same number of people, so that $t_j/T = t_i/T$ for any i and j. We introduce a variable $q \in [0, 1]$, which will determine the population shares in each of the J regions of the high polarization country, in the following way:

$$n_{1j} = 0.5 + 0.5q^j \text{ if } j \leq J/2 \tag{A1}$$

$$n_{1j} = 0.5 - 0.5q^{(j-J/2)} \text{ if } j > J/2. \tag{A2}$$

As constructed, the population share of dissatisfied group $i = 1$ will gradually be increasing with q in the first $J/2$ regions, and decreasing in the remaining $J/2$ regions.[14] After some algebraic manipulation we can express segregation only as a function of q and J (equation A3), i.e. via

[14]Note that there are infinitely many ways to restrict the way in which we construct the regional population shares n_{1j} and n_{2j}. Important restrictions are that regional population shares must be between 0 and 1, that in each region they add up to 1, and that they add up so that country wide population shares are as defined. An "easy" alternative specification would have been $n_{ij} = 0.5 + 0.5q^{j/x}$, where x is any positive number.

variable q, we have built an injective map from population shares to segregation. At maximum polarization, obtained when $n_1 = n_2 = 0.5$, and given the regional population shares in equations A1 and A2, we get the following:

$$\text{Segregation}_{\text{Pol}=1} = \sum_{j=1}^{J/2} \frac{2}{J} q^{2j} \tag{A3}$$

It is easy to see in equation A3 that when $q = 0$ segregation is equal to zero, when $q = 1$ segregation is equal to one, and in between segregation increases with q.

Next, let us consider a case in which the country-wide population share of the dissatisfied group is, for example 5%, so polarization is low, at 0.19. In this case $n_1 = 5\%$ and $n_2 = 95\%$. Let us construct the regional population shares in the following way:

$$n_{1j} = 0.05 + 0.95q^j \text{ if } j \leq 0.05J \tag{A4}$$

$$n_{1j} = 0.05 - 0.05q^{(\text{floor}\,\frac{(j-1)}{0.05J})} \text{ if } j > 0.05J \tag{A5}$$

Substituting the population shares from equations A4 and A5 into the formula for segregation given in equation 5, we can can again derive segregation as a function of only J and q. However, in the low polarization country the final expression for segregation would be more cumbersome than in equation A3, and we omit it from the text.

The next step is to compute the first-period total effort in conflict, and first-period attacks. As shown in the model in Section 2, both of these are functions of regional population shares, and now given the injective map from population shares to segregation, we can express first-period conflict outcomes as a function of segregation, shown in Figure A12. Note that in the first period, conflict occurs only in regions where the population share of the dissatisfied group exceeds 61.8%, so given the population shares defined in equations A1 and A2, we can expect conflict only in the first $J/2$ regions in the high polarization country and in the first $0.05J$ regions in the low polarization country. Total effort expended in conflict is the sum of efforts of both groups in each region, and given that $p_{ij} = n_{ij}$ in the first period, we have that $E = \sqrt{n_{1j}(1 - n_{1j})}$ in the first period. Hence, the first-period effort expended in the whole country is given in equations A6 and A7 for the high and low polarization country, respectively.

$$E_{\text{period}=1,\text{ Pol}=1} = \sum_{j=1}^{J/2} \mathbf{1}_{\mathbf{n_{ij}}>\mathbf{0.618}} \sqrt{(0.5 - 0.5q^j)(0.5 + 0.5q^j)} \tag{A6}$$

$$E_{\text{period}=1,\text{ Pol}=0.19} = \sum_{j=1}^{0.05J} \mathbf{1}_{\mathbf{n_{ij}}>\mathbf{0.618}} \sqrt{(0.95 - 0.95q^j)(0.05 + 0.95q^j)} \tag{A7}$$

Using equations A6 and A7 it can be shown that for a given segregation level (i.e. given q) total effort in the first period in the high polarization country is higher than in the low polarization country. Plotting these functions in Figure A12 for different values of q – and hence different levels of segregation – we see that total effort reaches its peak around segregation of 0.3 for the high polarization country, and around segregation of 0.6 for the low polarization country. Therefore, the manner in which segregation is related to conflict is different in high and low polarization countries. We investigate this claim more systematically in the paper through simulations.

A.6 Which dataset to use for empirical analysis of segregation?

It is difficult to study the effect of segregation on conflict using real-world data. The key independent variable to test our theoretical claims is the index of segregation, for which we require very fine-grained data on ethnic composition. To construct segregation indices at the country level, we require ethnic group population shares at the sub-national level, ideally at a low administrative level such as municipalities or settlements. This is because people often segregate at the level of villages, settlements, or neighborhoods, and if population shares are taken over larger units segregation might seem artificially low.

Cross-national datasets widely used in studies of conflict, such as GeoEPR (Wucherpfennig et al. 2011) and PRIO-GRID (Tollefsen et al. 2012) contain a wealth of local-level information on the presence of different ethnic groups, but they do not contain subnational population shares. Hence, they are unsuitable for our study, because they do not allow us to calculate the index of segregation. Other prominent cross-national datasets such as Minorities at Risk (Gurr et al. 2009) or the dataset introduced in Weidmann et al. (2010) are similarly unsuitable for testing our theory. While they have information on related concepts of group concentration, they are insufficiently detailed to reveal the precise level of group mixing at a subnational level, needed to calculate the index of segregation.

Alesina and Zhuravskaya (2011) are the first to provide cross-country segregation data with individual group population shares at the subnational level. However, these data suffer from several important shortcomings in the context of our study. First, the ethnic composition data are drawn from census and other data sources only around the year 2000 (p. 1879). It would be unreasonable to use ethnic composition data from the 2000s to predict conflict outcomes in decades removed from the same period. This restricts the number of conflict episodes that could be studied. Second, conflict is quite rare in the set of countries for which segregation data are available. Finally, the level of disaggregation at the subnational level is often insufficiently low and quite variable across countries. For example, the number of regions in Macedonia, with a population of about 2 million, is 75, whereas the number of regions in Pakistan, with a population of about 180 million, is six.

A.7 Summary Statistics of Indian and Bosnian data

Summary statistics for the 314 Indian districts in 1981 and 389 districts in 1991, taken from respective decennial Censuses. Across Indian districts, we observe the full range of polarization, from 0 to 1, but the average polarization tends to be rather low at about 0.34 and 0.38 (in 1981 and 1991), which corresponds to the average Muslim share of the district population of about 10%. Segregation ranges from the minimum of 0 to approximately 0.50 and 0.58, with median segregation of 0.03 and 0.04 in 1981 and 1991, respectively. This is a significantly smaller range than in our simulated data. The main reason is that it could be argued that even tehsils are *too large* a unit. If information on Hindu and Muslim population were available at the sub-tehsil level, most likely we would have been able to obtain a greater range for segregation. This is because certain blocks within a city, or even a whole village might be either Hindu or Muslim, but the entire tehsil is highly unlikely to be composed of only one group; rather, it has a mix of the two groups. However, the census does not release ethnic composition at lower levels of administration, deeming this potentially sensitive information.

Summary statistics for Bosnia are given in Table A3 below.

Figures A13 and A14 show the distribution of segregation and polarization in India and Bosnia, respectively.

A.8 Robustness Tests

Due to space constraints in the main text, we present the main robustness tables for India and Bosnia in Tables A4 and A5 below. The results from these tables are discussed in the main text.

There are several additional potential concerns with regards to the empirical results that we address here. We begin with results based on Indian data. In some parts of India, groups other than Hindus and Muslims are present, so we check that our results hold if we restrict the sample to districts where the joint population share of Hindus and Muslims is greater than 70% (column 1 in Table A6 for the 1980s and Table A7 for the 1990s, respectively) and greater than 90% (column 2 in each table). Moreover, a careful reader might notice from Table A2 that there are districts in our sample with only one tehsil (the mean number of tehsils is 9). In column 3 in each table we show that our results hold if we restrict our sample to districts with more than two tehsils.

Finally, minimum segregation need not necessarily mean that the Muslims are evenly spread throughout the district. It may also arise in districts with an overwhelming majority of Hindus and a negligible minority of Muslims. This is mainly an artifact of the fact that tehsils still represent a relatively high level of disaggregation. We confirm in column 4 in Table A6 for the 1980s and Table A7 for the 1990s, respectively, that the results are broadly unchanged when we keep only districts with more than two percent Muslims.

In Table A8, we conduct equivalent sensitivity tests for results based on Bosnian data. Our estimates are substantively unchanged if we restrict the analysis to municipalities where the joint share of the two largest groups is greater than 70% (column 1) or 90% (column 2), or to municipalities comprised of five or more settlements (column 3).[15]

References

Alesina, Alberto and Ekaterina Zhuravskaya, "Segregation and the Quality of Government in a Cross-Section of Countries," *American Economic Review*, 2011, *101(5)*, 1872–1911.

Gurr, Ted Robert, Monty G Marshall, and Christian Davenport, "Minorities at Risk Project," *College Park, MD: Center for International Development and Conflict Management*, 2009.

Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug, "PRIO-GRID: A Unified Spatial Data Structure," *Journal of Peace Research*, 2012, *49* (2), 363–374.

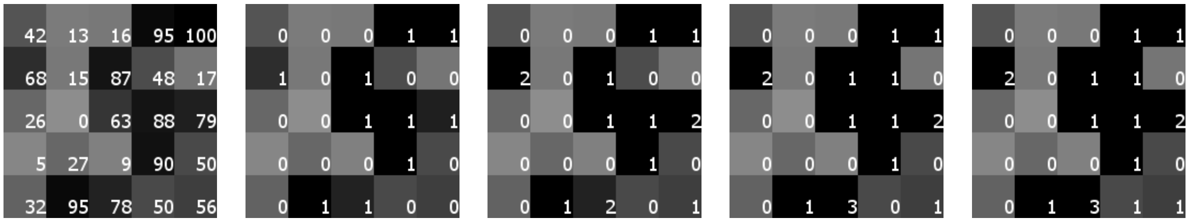
Weidmann, Nils B, Jan Ketil Rød, and Lars-Erik Cederman, "Representing ethnic groups in space: A new dataset," *Journal of Peace Research*, 2010, *47* (4), 491–499.

Wucherpfennig, Julian, Nils B. Weidmann, Luc Girardin, Lars-Erik Cederman, and Andreas Wimmer, "Politically Relevant Ethnic Groups across Space and Time: Introducing the GeoEPR Dataset," *Conflict Management and Peace Science*, 2011, *28* (5), 423–437.

[15] We do not conduct the sensitivity test analogous to that for India presented in column 4 in Tables A6 and A7. The lowest shares of the second largest group in Bosnian municipalities are mostly larger than the minimum share of Muslims in India.

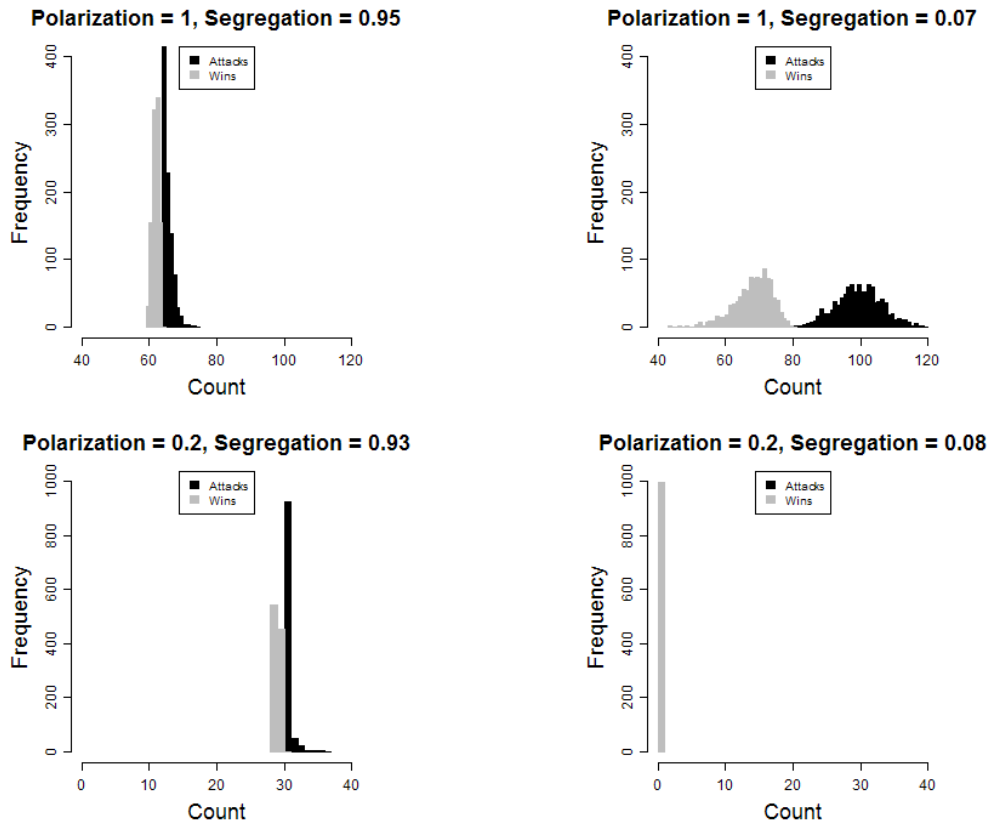
Figure A1: One example of the spread of conflict

(a) Regional shares of the dissatisfied group (b) Period 1 attacks (c) Period 2 attacks (d) Period 3 attacks (e) Period 4 attacks



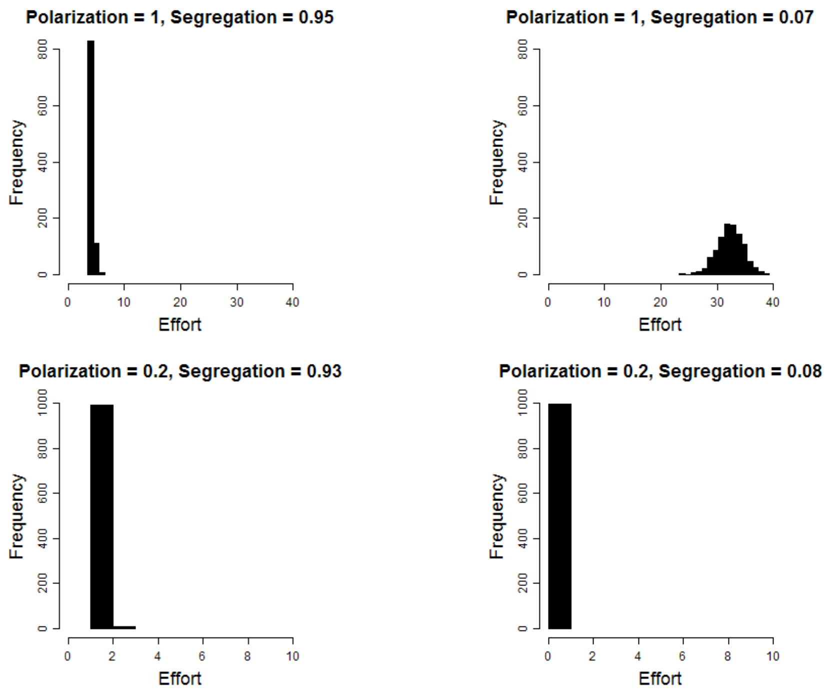
Note: Figure A1a gives an example of hypothetical country with 25 regions, and regional population shares of the dissatisfied ethnic group as indicated (shades of gray correspond to the share of dissatisfied ethnic group in each region). Figures A1b to A1e depict one possible history of conflict in this country. The first period attacks (Figure A1b) must always be the same in any run of the model, because they only depend on population shares. Attacks in later periods depend on regional population shares and battle successes of co-ethnic neighbors. When a region is captured, it turns black, and the total number of attacks in that region remains the same in future periods.

Figure A2: Distribution of total attacks and victories for countries from Figure 1 in the main text



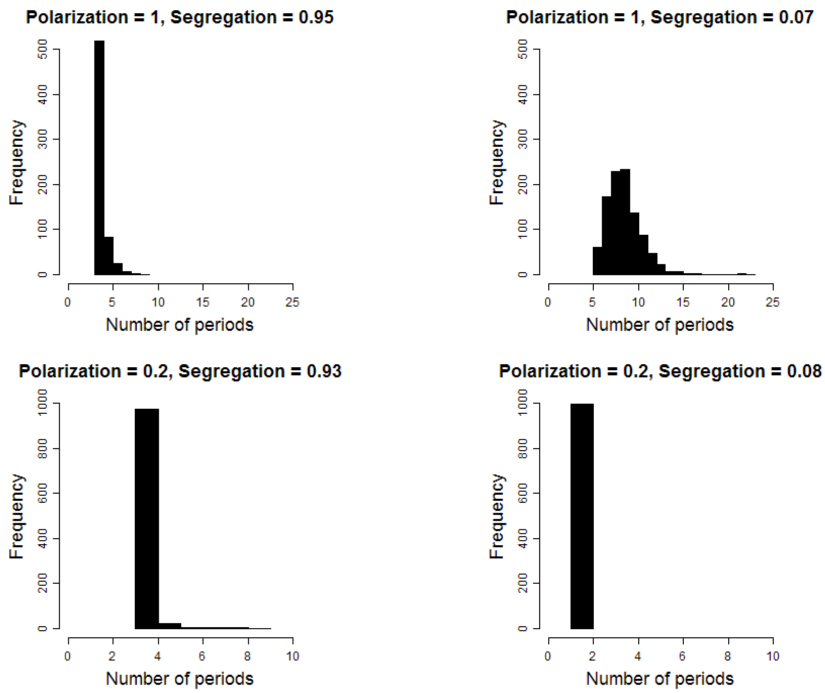
Note: For summary statistics corresponding to this distribution see Table A1.

Figure A3: Distribution of total effort for countries from Figure 1 in the main text



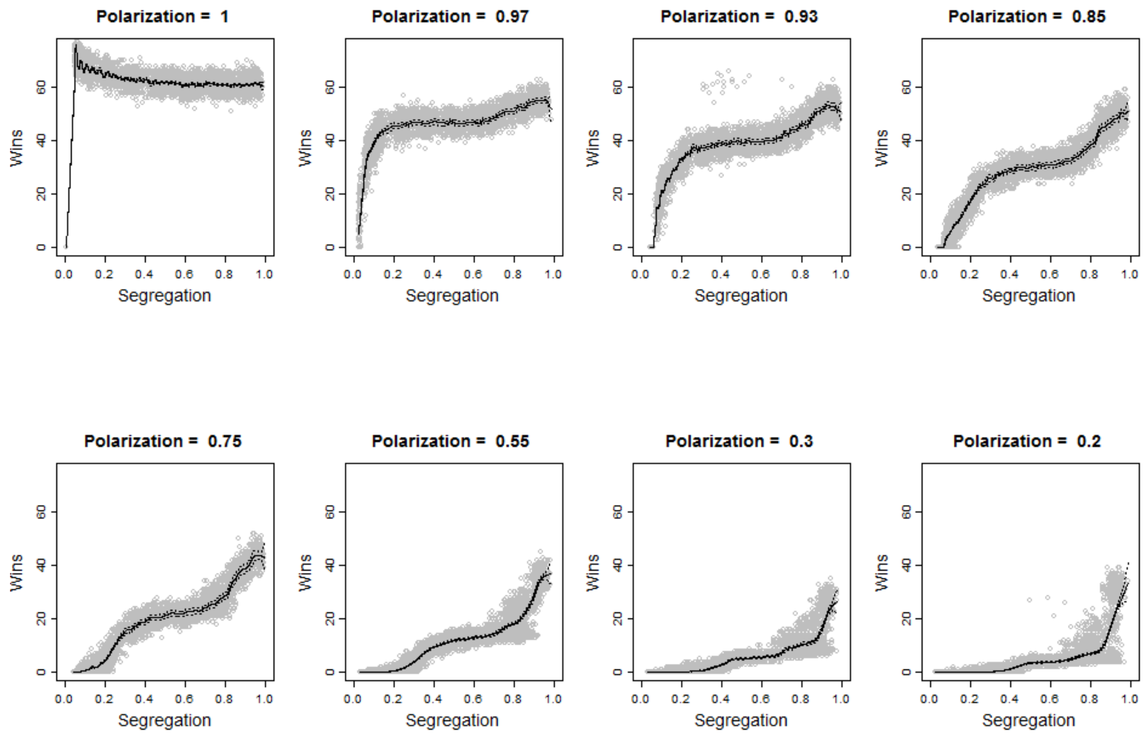
Note: For summary statistics corresponding to this distribution see Table A1.

Figure A4: Distribution of conflict duration for countries from Figure 1 in the main text



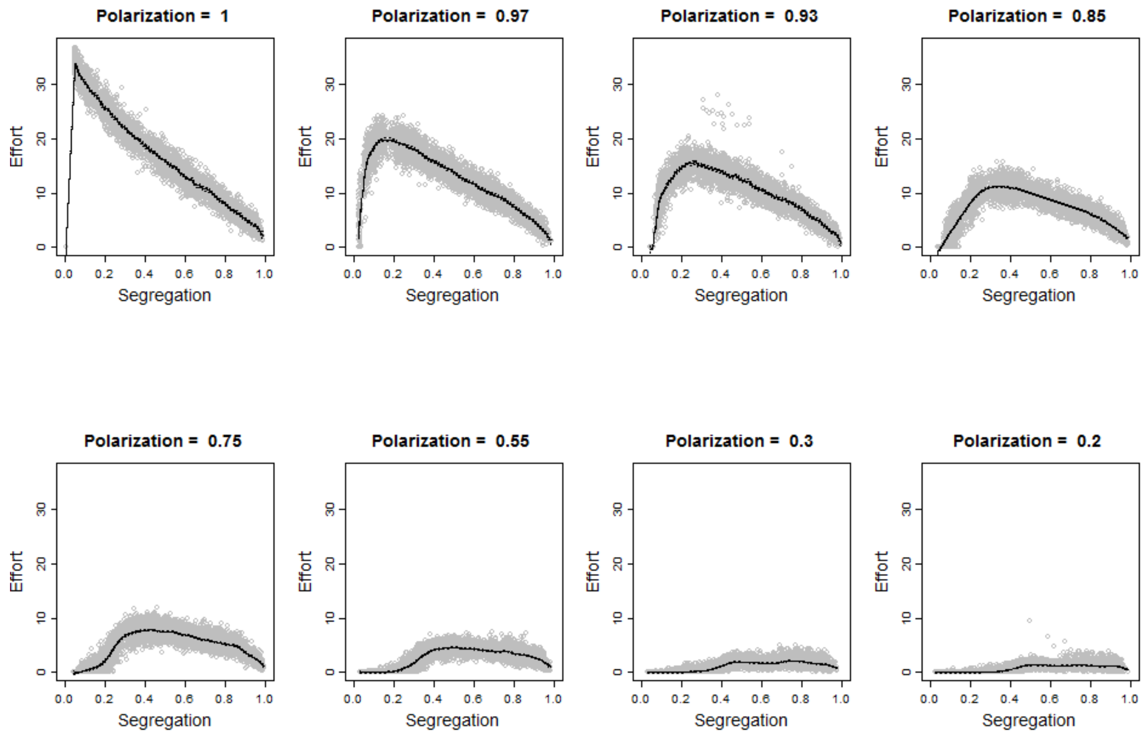
Note: For summary statistics corresponding to this distribution see Table A1.

Figure A5: Relationship between ethnic segregation and the number of victories, for eight levels of polarization



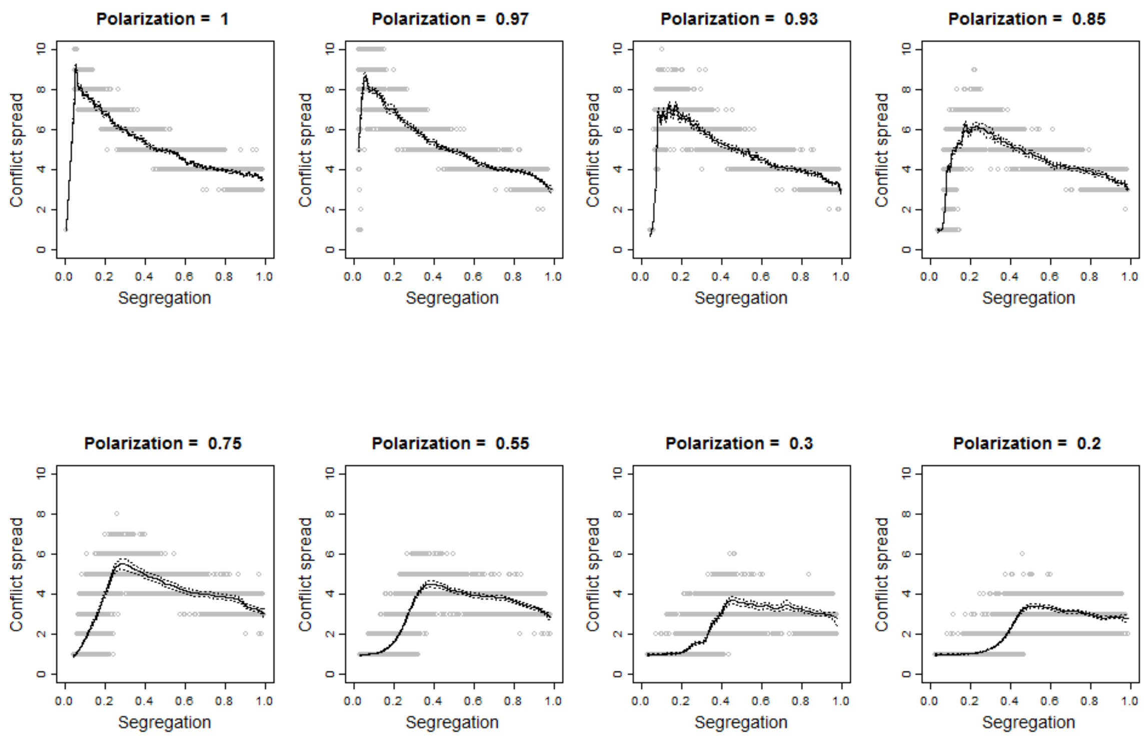
Note: Grey dots are outcomes of simulations. Black lines represent estimates from a bivariate generalized additive model (GAM) of victories on segregation, for each polarization level separately. Approximate 95 percent confidence intervals are also plotted.

Figure A6: Relationship between ethnic segregation and effort, for eight levels of polarization



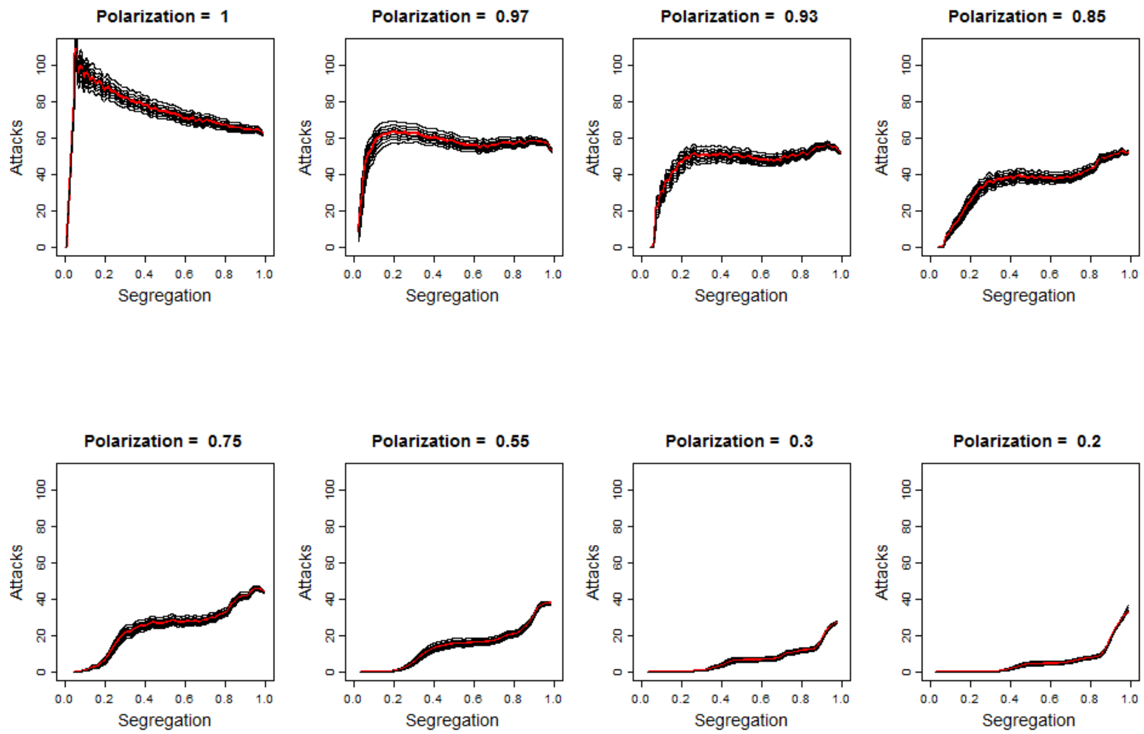
Note: Grey dots are outcomes of simulations. Black lines represent estimates from a bivariate generalized additive model (GAM) of effort on segregation, for each polarization level separately. Approximate 95 percent confidence intervals are also plotted.

Figure A7: Relationship between ethnic segregation and conflict duration, for eight levels of polarization



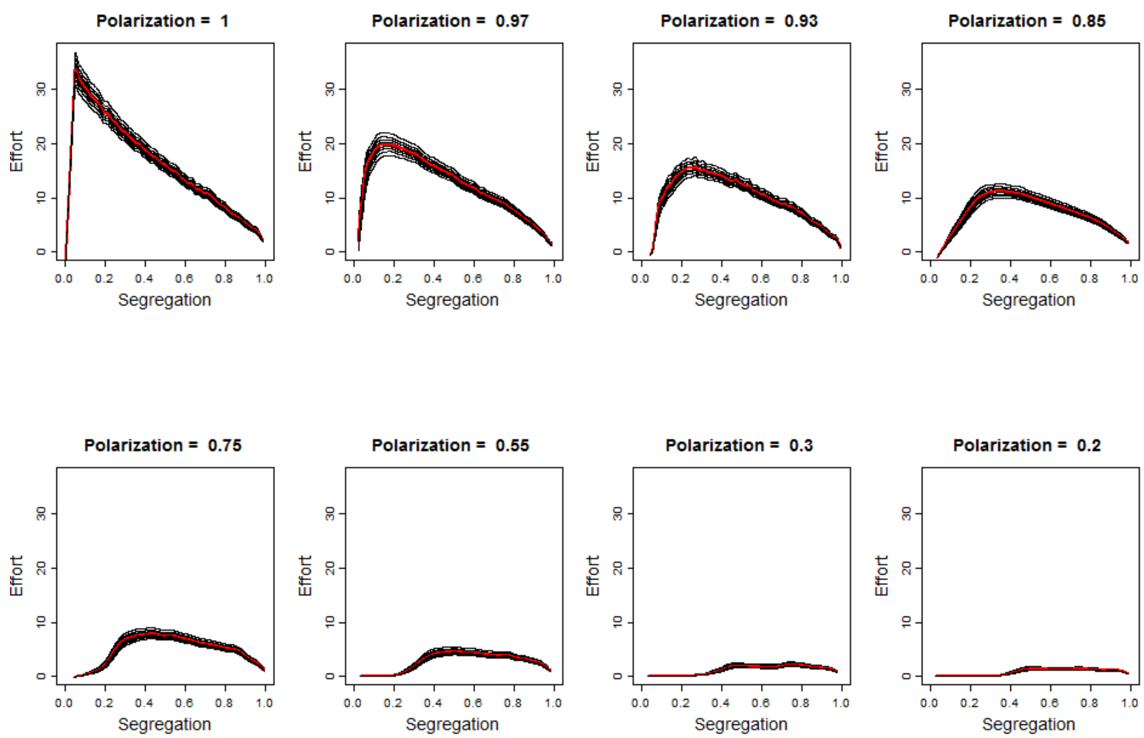
Note: Grey dots are outcomes of simulations. Black lines represent estimates from a bivariate generalized additive model (GAM) of conflict duration on segregation, for each polarization level separately. Approximate 95 percent confidence intervals are also plotted.

Figure A8: Relationship between ethnic segregation and each decile of attacks, across eight levels of polarization



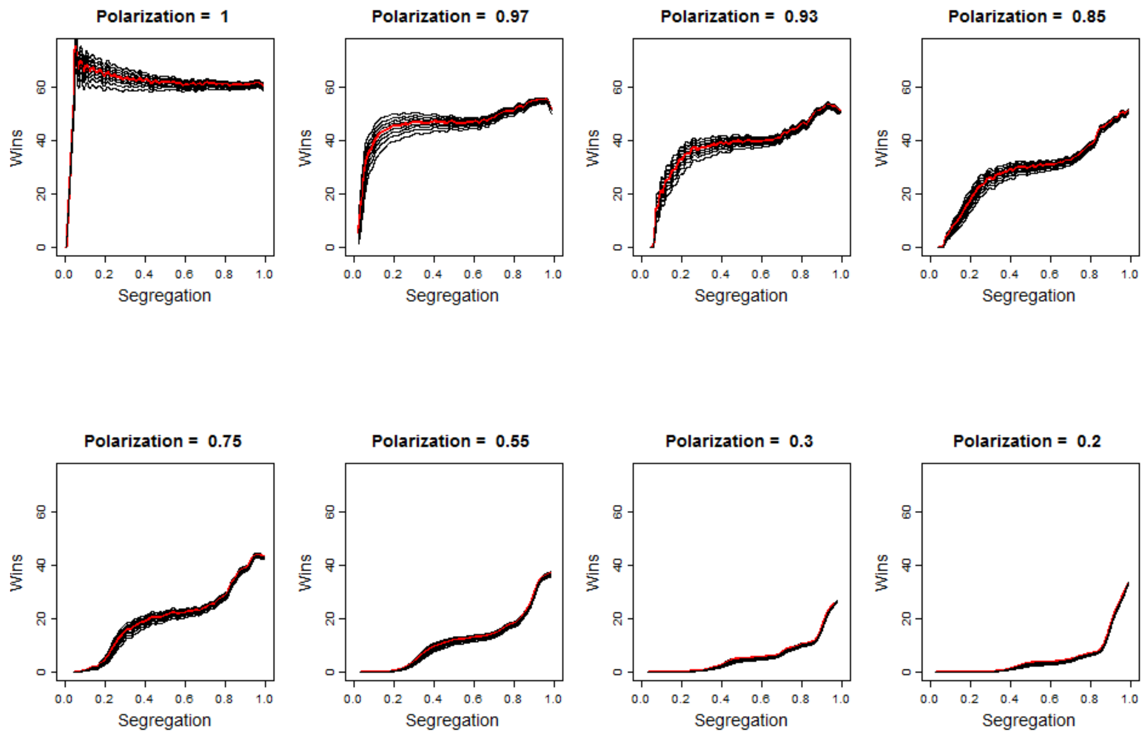
Note: Lines represent estimates from a generalized additive model (GAM) using thin plate regression splines. Number of attacks is assumed to follow the negative binomial distribution. The red line represents estimates on median outcomes also shown in Figure 2, whereas dark lines are estimates on the remaining 9 deciles.

Figure A9: Relationship between ethnic segregation and each decile of total effort, across eight levels of polarization



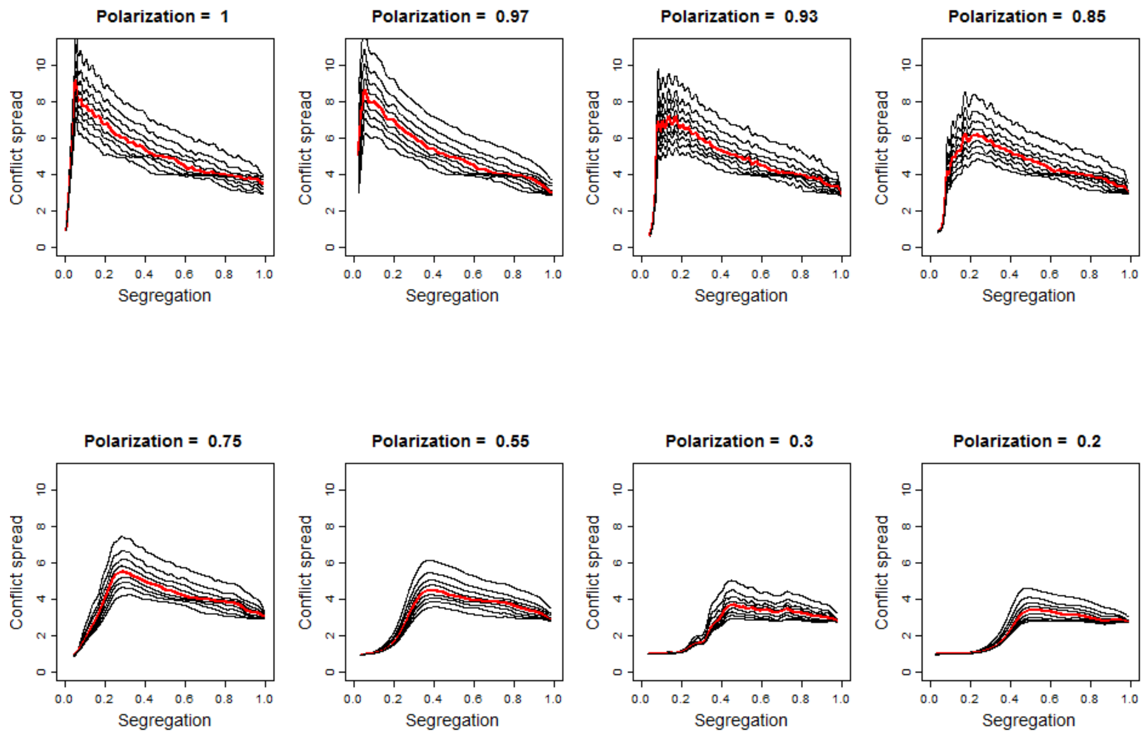
Note: Lines represent estimates from a generalized additive model (GAM) using thin plate regression splines. Effort is assumed to follow the standard Gaussian distribution. The red line represents estimates on median outcomes also shown in Figure A6, whereas dark lines are estimates on the remaining 9 deciles.

Figure A10: Relationship between ethnic segregation and each decile of victories, across eight levels of polarization



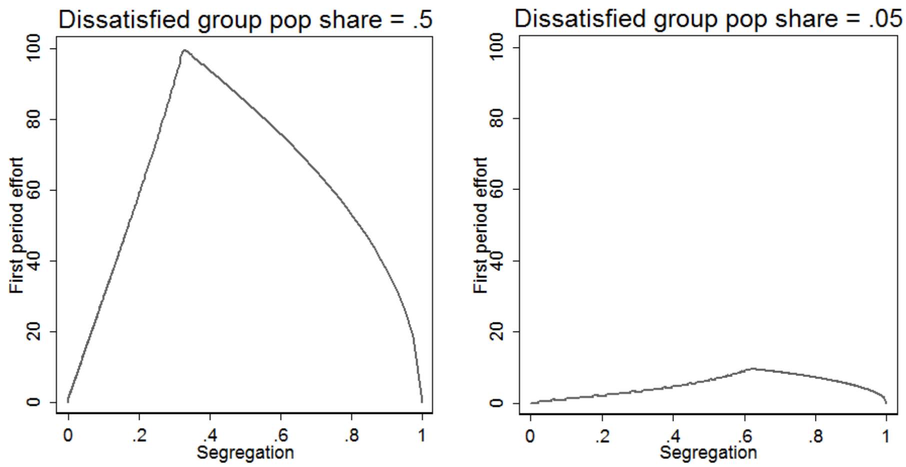
Note: Lines represent estimates from a generalized additive model (GAM) using thin plate regression splines. Number of victories is assumed to follow the negative binomial distribution. The red line represents estimates on median outcomes also shown in Figure A5, whereas dark lines are estimates on the remaining 9 deciles.

Figure A11: Relationship between ethnic segregation and each decile of conflict duration, across eight levels of polarization



Note: Lines represent estimates from a generalized additive model (GAM) using thin plate regression splines. Duration of conflict is assumed to follow the Gamma distribution. The red line represents estimates on median outcomes also shown in Figure A7, whereas dark lines are estimates on the remaining 9 deciles.

Figure A12: Special case – segregation and first-period effort



Note: First period effort jumps up each time that a new region reaches 61.8% of the dissatisfied group that is required for attack. As the population share of the dissatisfied group increases in the regions experiencing first period conflict, total effort expended decreases, until another region reaches the threshold of 61.8%. After all the regions dominated by the dissatisfied group are past the threshold for attacks, further increases in regional population shares, and thus increases in segregation, are associated with decreasing total effort. The figure is based on calculations done on 500 regions.

Figure A13: Distribution of segregation and polarization in Indian districts

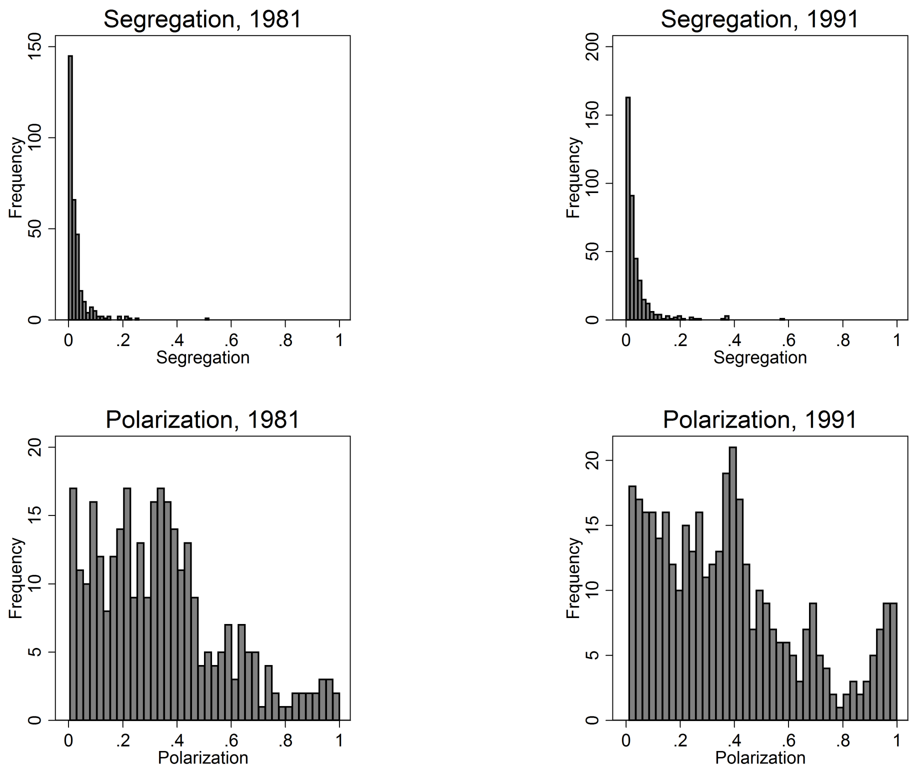


Figure A14: Distribution of segregation and polarization in Bosnian municipalities

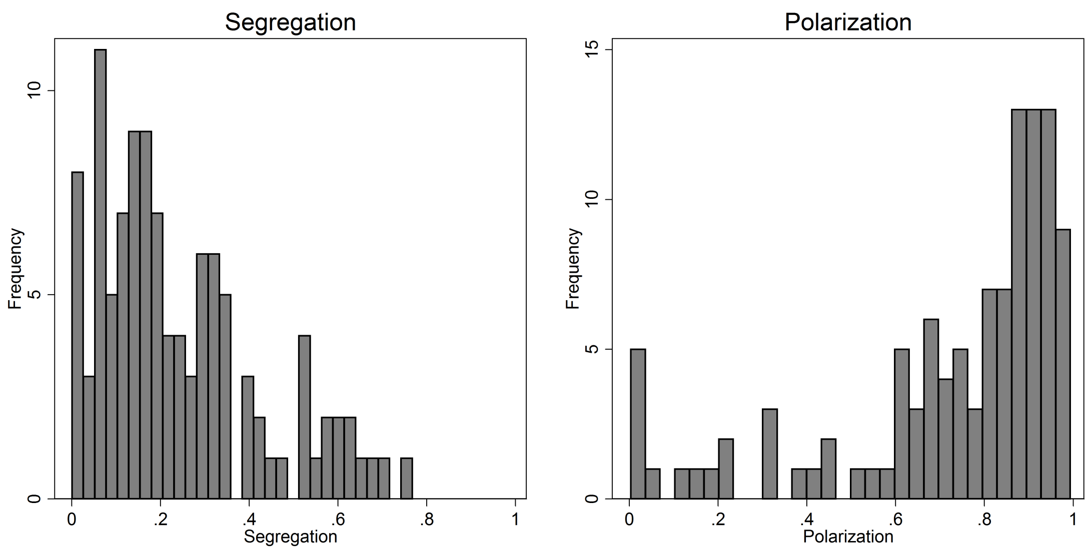


Table A1: Summary statistics for Figures A2, A3, and A4

	1st period Attacks	Total Attacks	Victories	Effort	Duration
Polarization = 1, Segregation = 0.95					
Min	59	64	59	3.48	2
Median	59	65	62	3.77	3
Max	59	74	64	6.19	7
Polarization = 1, Segregation = 0.07					
Min	20	72	43	23.33	4
Median	20	100	69	32.38	8
Max	20	119	80	38.69	21
Polarization = 0.20, Segregation = 0.93					
Min	28	30	28	1.00	2
Median	28	30	29	1.00	2
Max	28	36	30	2.95	7
Polarization = .20, Segregation = 0.08					
Min	0	0	0	0.00	0
Median	0	0	0	0.00	0
Max	0	0	0	0.00	0

Note: Summary statistics are based on 1,000 repetitions of the conflict model for each country in Figures 1 in the main text. The segregation levels were obtained as explained in the note to Figure 1 in the main text.

Table A2: Summary Statistics, India

	Obs	Mean	Median	St. Dev.	Min	Max
1980s						
Polarization	314	0.34	0.32	0.23	0.00	1.00
Segregation	314	0.03	0.02	0.05	0.00	0.52
Hindu share	314	0.84	0.88	0.14	0.21	1.00
Muslim share	314	0.10	0.08	0.09	0.00	0.66
Joint share	314	0.94	0.98	0.13	0.21	1.09
Income scale	314	-0.00	-0.18	0.98	-1.68	4.65
Population	314	1922776	1802623	1159459	32100	10739439
Number of tehsils	314	9	6	7	1	41
Number of riots	314	1	0	4	0	50
1990s						
Polarization	389	0.38	0.34	0.27	0.01	1.00
Segregation	389	0.04	0.02	0.06	0.00	0.58
Hindu share	389	0.81	0.87	0.17	0.04	0.99
Muslim share	389	0.12	0.09	0.13	0.00	0.95
Income scale	389	-0.00	-0.23	0.99	-1.45	4.77
Population	389	2100904	1904278	1223346	33224	8784445
Number of tehsils	389	11	7	11	1	66
Number of riots	389	1	0	2	0	25

Note: In the 1980s, the income factor scale contains the share of rural population, working male population, farmers, miners, literates, those enrolled in primary education, graduate education, and the share of migrants. All shares are expressed in terms of total district population. In 1991, in addition to the variables already mentioned, the income factor scale contains the share of households owning houses built of mud/grass/wood, tile/slate/shingle, metal, cement sheet, brick, stone, and concrete, as well as the share of households with access to electricity, closed toilet, and water. The factor scales exhibit construct validity: the first factor is significantly stronger than the remaining ones, suggesting that the variables capture well a single latent variable. The inter-item correlation among the variables is 0.76 in the 1980s and 0.73 in the 1990s data.

Table A3: Summary Statistics, Bosnia

	Obs	Mean	Median	St. Dev.	Min	Max
Polarization	109	0.73	0.83	0.27	0.00	0.99
Segregation	109	0.24	0.19	0.18	0.00	0.77
Serbs share	109	0.35	0.28	0.27	0.00	0.97
Muslim share	109	0.39	0.40	0.24	0.00	0.97
Croat share	109	0.20	0.07	0.27	0.00	0.99
Joint share (two largest)	109	0.89	0.90	0.09	0.63	0.99
Municipal Income, 1991 (logged)	109	11.84	11.72	0.95	9.92	14.29
Population	109	40042	30624	33079	4162	195139
Number of settlements	109	20	17	14	3	67
Number of attacks	109	15	8	16	0	76

Table A4: Robustness of results to additional controls, India

	1980s			1990s		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Polarization index	3.50***	3.34**	4.24***	3.50***	3.43***	3.33***
	(1.03)	(1.36)	(0.97)	(0.72)	(0.78)	(0.92)
Segregation index	18.38**	16.92*	16.81**	20.33***	23.65**	29.47*
	(9.09)	(8.87)	(7.31)	(7.64)	(11.00)	(15.66)
Polarization × Segregation	-28.90**	-27.42**	-26.28**	-31.18***	-34.86***	-40.56**
	(12.15)	(12.74)	(11.42)	(9.00)	(12.86)	(18.96)
Income scale	0.62***	0.38		0.57***	0.85**	
	(0.21)	(0.34)		(0.20)	(0.36)	
Constant	-1.49***	-2.27	5.46	-2.12***	-1.13	5.87*
	(0.55)	(1.75)	(3.83)	(0.41)	(1.28)	(3.14)
Elevation 250-500m (% of district)		-0.47	-0.13		-0.80	-0.96
		(0.44)	(0.54)		(0.66)	(0.74)
Elevation 500-1000m (% of district)		0.79	0.47		0.51	0.12
		(0.56)	(0.40)		(0.50)	(0.69)
Elevation > 1000m (% of district)		-24.46**	-12.06***		-12.03**	-9.82**
		(10.54)	(2.90)		(5.18)	(3.91)
Literacy rate		2.74	2.11*		-1.85	0.99
		(4.10)	(1.22)		(2.73)	(1.62)
Rural population (% of total)			-1.84***			-3.79***
			(0.64)			(0.82)
Miners (% of total)			-44.91			-79.99**
			(37.28)			(37.09)
Migrants (% of total)			6.79***			2.32
			(2.46)			(3.30)
Poverty gap			-1.45			1.43
			(2.86)			(2.57)
Gini coefficient			-18.41***			0.76
			(4.12)			(3.55)
Rainfall (cm)			-0.00***			0.00
			(0.00)			(0.00)
Caste fractionalization index			-2.48			-6.57*
			(3.47)			(3.37)
Constant	0.81***	0.72**	0.03	1.01***	0.84***	0.90***
	(0.28)	(0.29)	(0.35)	(0.23)	(0.20)	(0.20)
Observations	314	306	262	389	306	255

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the number of riots in a district. Segregation and polarization are calculated based on the 1981 and 1991 census data. We use only the population of Hindus and Muslims, discarding other ethnic groups, and recalculating the ethnic shares of Hindus and Muslims accordingly. In the 1980s, the income factor scale contains the share of rural population, working male population, farmers, miners, literates, those enrolled in primary education, graduate education, and the share of migrants. All shares are expressed in terms of total district population. In 1991, in addition to the variables already mentioned, the income factor scale contains the share of households owning houses built of mud/grass/wood, tile/slate/shingle, metal, cement sheet, brick, stone, and concrete, as well as the share of households with access to electricity, closed toilet, and water. Standard errors (in parentheses) are clustered by state.

Table A5: Robustness of results to additional controls, Bosnia

	Model 1	Model 2	Model 3	Model 4
Polarization index	2.63***	2.67***	2.29***	2.35***
	(0.72)	(0.70)	(0.73)	(0.68)
Segreation index	8.68**	8.42*	7.55*	7.20*
	(4.30)	(4.32)	(4.11)	(3.96)
Polarization × Segregation	-10.40**	-10.26**	-9.14*	-8.27*
	(4.99)	(4.97)	(4.76)	(4.65)
Municipal income (log), 1991	0.17	0.16	0.12	0.09
	(0.12)	(0.12)	(0.13)	(0.13)
Constant	-1.42	-1.28	-0.55	-0.79
	(1.44)	(1.58)	(1.71)	(1.74)
Mean elevation		-0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)
Major road			0.41	0.57**
			(0.27)	(0.25)
Border dummy				0.55*
				(0.30)
Constant	0.35**	0.35**	0.33**	0.30*
	(0.16)	(0.16)	(0.16)	(0.16)
Observations	109	109	109	109

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: The dependent variable is the total number of attacks in a municipality. Segregation and polarization are calculated based on the 1991 census data including Muslims, Serbs and Croats. We use only the population of the three groups, discarding other ethnic groups, and recalculating the ethnic shares accordingly. Robust standard errors in parentheses.

Table A6: Additional robustness tests, India 1980s

	(1) Hindu/Muslim > .7	(2) Hindu/Muslim > .9	(3) > 2 Tehsils	(4) Muslim > .02
Polarization	3.10***	2.83***	3.49***	2.28**
	(1.00)	(0.97)	(1.09)	(1.07)
Segregation	16.56*	17.90*	16.39*	11.83
	(9.20)	(9.56)	(8.93)	(9.73)
Pol. × Segregation	-25.25**	-25.38**	-26.69**	-18.67
	(12.08)	(12.33)	(11.84)	(12.50)
Income scale	0.68***	0.77***	0.67***	0.62***
	(0.21)	(0.22)	(0.24)	(0.21)
Observations	297	268	299	273

Note: The dependent variable is the number of riots in a district. Segregation and polarization are calculated based on the 1981 census data. The income factor scale contains the share of rural population, working male population, farmers, miners, literates, those enrolled in primary education, graduate education, and the share of migrants. All shares are expressed in terms of total district population. Standard errors (in parentheses) are clustered by state.

Table A7: Additional robustness tests, India 1990s

	(1) Hindu/Muslim > .7	(2) Hindu/Muslim > .9	(3) > 2 Tehsils	(4) Muslim > .02
Polarization	3.18***	3.11***	3.42***	2.45***
	(0.71)	(0.64)	(0.73)	(0.79)
Segregation	18.40**	18.72**	19.09**	12.69**
	(7.60)	(8.30)	(7.42)	(6.26)
Pol. × Segregation	-28.22***	-28.34***	-29.74***	-20.75***
	(8.51)	(9.02)	(8.81)	(7.53)
Income scale	0.61***	0.68***	0.63***	0.56***
	(0.19)	(0.18)	(0.21)	(0.19)
Observations	367	329	375	336

Note: The dependent variable is the number of riots in a district. Segregation and polarization are calculated based on the 1991 census data. The income factor scale contains the share of rural population, working male population, farmers, miners, literates, those enrolled in primary education, graduate education, the share of migrants, the share of households owning houses built of mud/grass/wood, tile/slate/shingle, metal, cement sheet, brick, stone, and concrete, as well as the share of households with access to electricity, closed toilet, and water. All shares are expressed in terms of total district population. Standard errors (in parentheses) are clustered by state.

Table A8: Additional robustness tests, Bosnia

	(1) Joint Share > .7	(2) Joint Share > .9	(3) > 4 Subregions
Polarization	2.80***	3.62***	2.50***
	(0.71)	(0.65)	(0.78)
Segregation	9.24**	17.95***	8.14*
	(4.11)	(4.98)	(4.38)
Pol. × Segregation	-11.54**	-23.39***	-9.77*
	(4.78)	(5.83)	(5.09)
Municipal Income, 1991 (logged)	0.14	0.29	0.16
	(0.13)	(0.18)	(0.13)
Observations	103	57	105

Note: The dependent variable is the total number of attacks in a municipality. Segregation and polarization are calculated based on the 1991 census data including Muslims, Serbs and Croats. Robust standard errors in parentheses.