# SEGREGATION AND SOCIAL CONFLICT: AN EMPIRICAL ANALYSIS

**Alejandro Corvalán**

**and**

**Miguel Vargas**

**This draft: April 2013**
**Preliminary: not for citation**

## ABSTRACT

In this paper, we empirically investigate the relationship between ethnic segregation and social conflict. We believe that segregation can increase the collective articulation within groups and the difference between preferences, which can increase conflict intensity. Our focus is on ethnic segregation because we follow the idea that although conflicts can be economically motivated, they need other aspects to find their expression, such as religion, language or ethnicity. Using a different econometric specifications, we find robust evidence on the relationship between segregation and social conflict even after controlling for polarization, fractionalization and lagged conflict.

JEL Classification Nos.: D72, E21
Keywords: segregation, conflict, institutions, dynamic panel

## 1. Introduction

The aim of this work has been to investigate empirically if there is a relationship between segregation and social conflicts. Here we understand segregation at national level, therefore the segregation index used is a country level index. Specifically we used the country level ethnic segregation index calculated in Alesina and Zhuravskaya (2012). As dependent variables we have used two measures of conflict, the Peace Research Institute Oslo's measure (PRIO), a discrete variable which takes the value 1 if there is more than certain number of conflict related deaths in a period and 0 otherwise, and a continuous index of social conflict computed by the Cross-National Time-Series Data Archive (CNTS). This index of conflict is the weighted average of eight different manifestations of domestic conflict: Assassinations, General Strikes, Guerrilla Warfare, Major Government Crises, Purges, Riots, Revolutions, and Anti-government Demonstrations. Due to the characteristics of these two variables, we have used a probit model for the PRIO variable and a Tobit model for the CNTS one.

Since there is an endogeneity problem amidst social conflict and segregation, we have implemented an instrumental variable model as well. As instrument we used a predicted level of segregation computed in Alesina and Zhuravskaya (2012). This instrument considers, basically, that a country inhabited by a given ethnic group will be more segregated if this same ethnic groups lives in its neighboring countries.

As a final specification we run a panel data model. As segregation is a variable strongly correlated to countries fixed characteristics, such as institutional and geographical variables, following Acemoglu et al. (2009), we run a set of random effect regressions using geographical, institutional variables and segregation as proxies of countries fixed effects.

Our main result is that there is a positive and statiscally significant correlation between segregation and social conflict, even after controlling by traditional variables used in conflict literature, such as fractionalization and polarization indices, and lagged conflict.

We believe that a possible explanation of this positive correlation is that segregation can have effects on social conflict through two main factors: the collective action and the distance between

groups' preferences. Considering segregation as a particular institutional arrangement,[1] it will discriminate the spatial benefits in favor of the dominant social group, which, in turn, will generate tension between the different society groups, and this tension will be manifested in social conflict. The intensity of this conflict will depend on how effective the internal organization and cohesion within a group are and on how different the groups' interests are from each other. It is important to mention that this hypothesis have not been tested in the present investigation.

This ides has not been extensively studies as traditionally the literature on social conflict has established that social divisions are important for conflict. Popular measures of these divisions are fractionalization and polarization indices, which have played an important role in studying the determinants of social conflict (Collier and Hoeffler, 2004; Esteban and Ray, 2011, 2012; Fearon, 2003; Miguel et al., 2004; Montalvo and Reynal-Querol, 2005). Although these indices are important, they do not consider the spatial distribution of groups within a country or region, which is a different distributional aspect. For instance, a country could have a high level of fractionalization because there are several ethnic groups living in it but a low level of segregation because these groups are evenly distributed across the country, and vice versa.

Regarding the fact that we have used an ethnic segregation index instead of a socioeconomic one is based on the view that inequality is an important cause of conflict, however, as Esteban and Ray (2008) point out, conflicts could be motivated by economic aspects; nonetheless, they find their expression through the divisions caused by religion, ethnicity or national origins.

Following this argument, a high level of segregation would have as a consequence a low level of interaction between groups, because they do not share the same physical area. This lack of interaction could produce, amongst other things, two effects. First, it could increase the level of internal group cohesion, which would improve the mechanism for the articulation of collective actions. As Skirmuntt (2012) indicates, based on Giddens (1984) and Kirby (1989), the ethnic community model assumes that minorities tend to share the same kinds of problems and hence

[1] According to Knight (1992), a social institution is a set of rules that structure social interactions in particular ways, which are known by the members of the relevant community or society.

their interests are specific and diverge from those of the majority. Molina et al. (2002), following Alesina and LaFerrara (2000), Glaeser et al. (2002) and Portes (1998), point out that segregated neighborhoods can start collective action more easily than non-segregated neighborhoods by mobilizing ethnic or cultural ties, which is one of the transmission mechanisms linking segregation, individuals' outcomes and social capital. Consequently, when a large number of individuals with the same ethnic background live and share the experiences generated in the same geographical area, the probability of political involvement among them rises. This sense of belonging has been considered as one of the positive effects of segregation (see, for instance, Cheshire, 2007).

Second, when groups are segregated, the distance between their preferences will be greater than when members belonging to different groups have more chances of interaction. Conejeros and Vargas (2012) develop a model in which the lack of interaction between groups increases future prejudices and the distance between groups' preferences.

To the best of our knowledge, the unique previous effort related to this topic was undertaken by Matuszeski and Schneider (2006) and Novta and Klasnja (2012). The first work develops the diversity and clustering index, which measures the groups clustering in a country and the overall ethnic diversity. Using this index, the authors test the relationship between segregation and civil war. They find that civil war is more frequent in countries where inhabitants belonging to a given ethnic group live in a more clustered manner. Albeit interesting, this work does not include a set of controls that have been traditionally used in the literature, such as income, population, a country's political characteristics or other distributional measures, like polarization indices. Novta and Klasnja (2012) produce a theoretical model for understanding the relationship between segregation and polarization. After obtaining simulations from the model, the results are contrasted with Indian data on ethnic riots between Muslims and Hindus. They find that the effect of segregation on conflict depends on the level of ethnic polarization: with a high level of polarization an increase in segregation generates a less severe conflict and when polarization is low an increase in segregation makes the conflict more severe.

The difference between the current paper and the ones mentioned above is that our focus is on the intensity of conflict instead of the number of conflicts observed in a country and their duration. Besides, we try to exploit the temporal variation of some variables using panel data models instead of just cross-sectional regressions to research the relationship between ethnic segregation, at the regional level, and the intensity of social conflict. Finally, in our regressions, we control for fractionalization and polarization at the same time.

## 2. Data and Descriptive Statistics

### 2.1 Conflict

As a measure of conflict, our dependent variable, we consider the one based on the death toll provided by the PRIO data set (Gleditsch et al., 2002).[2] The PRIO data set constructs binary measures of conflict comparing the number of battle deaths per year with some given thresholds. The intensity of conflict is "low" if the death toll is higher than 25, "intermediate" if it ranges between 25 and 1000 and "high" (or war) if it exceeds 1000 deaths.[3] This measure has been used in a number of related works (Collier and Hoeffler, 2002; Doyle and Sambanis, 2003; Esteban and Ray, 2012; Fearon and Laitin, 2003).

Our baseline measure is low conflict. While we provide robustness checks for intermediate and high conflict, we mainly seek to explain the relation of segregation and small episodes of conflict. As the level of measurement – binary, ordinal, continuous – may affect the results (Benhabib et al., 2012), we check our results using a continuous index of social conflict as computed by the Cross-National Time-Series Data Archive (CNTS). The index is a weighted average over eight dimensions of internal conflict adopted from Rummel (1963). While the measure is continuous, we notice that it has substantial mass at the boundary in zero.

[2] This is a joint data set of the Uppsala Conflict Data Program (UCDP) and the Centre for the Study of Civil War at the International Peace Research Institute, Oslo (PRIO). It is available at http://www.prio.no/Data/.
[3] A natural concern is that those figures are not normalized by country size. We take care of the problem by controlling by population in all our exercises.

Given our data on episodes of conflict, there are three groups of conflict measures to be studied. First is the the onset or outbreak of conflict, which is related to the decision to trigger a conflict and can be defined as a year with conflict with a number of previous years without conflict (Strand, 2006). Some studies using this kind of measure are Collier and Hoeffler (1998), Esteban and Ray (2008) and Weidmann, (2009). Second, some studies are related to the duration of conflict, such as Collier et al. (2004) and Fearon and Laitin (2003). Finally, the intensity of conflict simply measures whether conflict is zero or one in a particular year. Studies relating the intensity of conflict and several ethnic factors are Collier and Hoeffler (1998), Esteban and Ray (2012), Fearon and Laitin (2003), Miguel et al. (2004), Montalvo and Reynal-Querol (2005) and Weidmann (2011). The factors that contribute to one dimension of conflict may not affect another dimension.[4] In this paper, we focus on the incidence of conflicts.

## 2.2 Segregation

For segregation we use Alesina and Zhuravskaya's (2011) computations. They construct an index that ranges from 0 to 1, where a value of 1 implies full segregation and a value of 0 the opposite. This index, conceptually, considers that for full segregation every ethnic group should live in separate regions within a country; meanwhile, for null segregation every region must have the same ethnic group composition as the country as a whole. This index of ethnic segregation was calculated for 97 countries using census information closest to the year 2000 as the main source of information. The second source of data they use, if census data are not available, is the statistics published by the national statistics agency of the countries. If neither of these two sources of information is available, they use the regionally representative Demographic and Health Surveys.

The index is a multi-group index; therefore, all the possible groups existing in a country are considered for these calculations, unlike traditional segregation indices, which take into account just two arbitrarily defined groups. Due to the latter, these authors propose a correction for the case of those groups, which are classified as "others," assuming that they are a composition of small ethnic groups with no segregation within them, which means that the subgroups of the

[4] Schneider and Woesehomeier (2006) show that the factors that contribute to the outbreak of a conflict do not coincide with the ones that keep feeding it.

"others" category are evenly distributed across all regions. The numbers of "others" subgroups are set equal to the number of people in "others" divided by the size of the smallest group identified.

Segregation data do not change over time. In the cross-countries analysis this is not an issue; nevertheless, in the panel data analysis it deserves further discussion. First of all, the literature about segregation, at least at the residential level, has established that this is a stable phenomenon in the sense that segregation is a stable equilibrium while integration is an unstable one; therefore, in the long run, we will always observe segregation. The main forerunner of this literature is Schelling (1971), who shows that even if everyone wishes to live in mixed communities, the sum of individuals' free choices will generate segregated communities. Later works by Krugman (1996) and Young (1998) demonstrate that mixed communities are stochastically unstable, which means that small random shocks to an integrated community will lead to a breakdown that will end up in full segregation.

In terms of conflict, if two groups are fighting for control of a territory, the final outcome can be the exchange of the dominant group in a given region and probably the eviction of the defeated group towards the land formerly occupied by the winning ethnic group, a situation that will leave the segregation more or less the same, at least in the dimensions measured by the Alesina and Zhuravskaya (2011) index.

*Instrument*

Social conflicts may leave to migrations either between regions within a country or between different countries. As segregation is defined by the place where people choose to live, which in turn depends on migration, an endogeneity problem could occur. In order to deal with this problem, we use the instrument developed by Alesina and Zhuravskaya (2011). This instrument is based on the idea that people belonging to a particular group gravitate towards the borders of countries with the same ethnic groups. With this in mind, the predicted location of members of each group in each country is computed. Based on these predictions, an index of predicted segregation is calculated, which is used as an instrument for observed segregation.

The intuition is as follows: if in a country there is a group that also lives in a neighboring country, then it is likely that this group will be located close to the border with the country with

the same ethnic group, like Basques in the north of Spain and the South of France. On the other hand, if an ethnic group is not present in any of the neighboring countries, it is probable that it will be evenly spread across the country. This could be the result of the historic formation of borders, the gradual spreading out across borders or the awkward drawing of borders that split groups into two adjoining countries.

The calculation of this index is carried out as follows. First, a home country is considered as having *K* neighbors. It is assumed that it has K hypothetical regions. A predicted distribution of people into these regions is constructed, assuming that the members of each group gravitate towards the regions that are closer to the borders of countries that have the same ethnic group. Finally, the predicted segregation index is obtained on the basis of the predicted location.

### 2.3 Other Data

Several other variables may affect both conflict and segregation. We include four sets of variables. The first set groups all the variables that change over time during the post-war period. All the rest remained fixed over time, as they intend to capture country-specific effects.

*Non-fixed Variables*

As conflict considers an absolute number of casualties, we control for country populations. Income also plays a major role in the estimation, since income and growth are strongly and negatively related to civil conflict (Collier and Hoffler, 1998; Miguel et al., 2004). The more complete data set for population and income comes from Maddison (2008).

Finally, political institutions may play a role in regulating and canalizing conflict. We consider the composite Polity2 index for democracy, which measures the difference between the Polity IV Democracy and Autocracy indices and ranges between -10 and 10. As for robustness, we check that our results are unchanged if we use Population and GDP data from PWT (7.0) and the binary Democracy–Dictatorship index from Alvarez et al. (2000).[5]

[5] Maddison covers 202 countries and PWT covers 189 countries. DD is an annual dummy variable that takes a value of 1 in the presence of an authoritarian regime.

*Geographical Variables*

Several geographical and endowment factors reinforce conflict. The typical variables used in the literature are the following: "mountains," measured as the percentage of mountainous terrain (Fearon and Laitin, 2003); "ncont" or non-contiguous states, referring to countries with territory holding at least 10,000 people and separated from the land area containing the capital (Fearon and Laitin, 2003); and a diamond dummy, which takes the value 1 if the country is rich in "oil" or produces (any positive quantity of) diamonds (Ross, 2012).

*Ethnic Variables: Fractionalization and Polarization*

The fractionalization index measures the probability that two people drawn at random will belong to different ethnic groups; meanwhile, polarization compares the ethnic groups' relative size. For instance, if two main ethnic groups within a region are similar in size, polarization is high, but if one of them is relatively small, polarization is low. For both the fractionalization and the polarization index we use Esteban and Ray's (2011) calculations. Equations (*) and (**) show the formal expressions for fractionalization and polarization, respectively.

$$(*) \qquad F = \sum_{i=1}^{m} n_i(1 - n_i)$$

$$(**) \qquad P = \sum_{i=1}^{m} \sum_{j=1}^{m} n_i^2 n_j d_{ij}$$

where ***n*** represents each group size, ***m*** the number of groups and ***d*** the distance between groups' preferences.

For the empirical computation of these indices, Esteban and Ray (2011) use James Fearon's data set, which is an update of the one in Fearon (2003). This data set identifies over 800 ethnic and ethno-religious groups in 160 countries.

Regarding preferences' distance over public goods, the linguistic distance is used as a proxy. Following the idea that languages spoken can be organized in a tree, which captures their genealogy, the similarity between two languages, $s$, is defined as the ratio of the number of common language branches to the maximum possible number, fifteen according to Ethnologe, and the language distance is defined as $k=1-s^d$, where $d=0.05$. Esteban and Ray (2011) present a discussion about why the latter should be the baseline value of $d$.

*Long-Run Variables: Latitude and Institutions*

Several studies show the existence of long-run effects of natural endowments and original institutions over the current growth rate and institutional arrangements. Acemoglu et al. (2009) indicate that suppressing those factors may generate an omitted variable bias in post-war panel estimations. In our case, some underlying long-run determinants may generate segregation and conflict during the whole period. To deal with that possibility, first we consider the tropic hypothesis (Landes, 1998; Sach and Wagner, 1997) and we add latitude as a major long-run control. Second, we control for historical institutions. Several variables have been proposed as proxies for them (Acemoglu et al., 2001; Easterly and Levine, 2003). Here we follow Acemoglu et al. (2009), who provide a set of variables aimed to approximate institutional country effects in the estimation of the modernization hypothesis. A high indigenous population density is a signal of greater extraction of resources and repression by the Europeans, so it predicts a worse institutional arrangement. As this measure is noisy, it is also convenient to consider a direct measure of institutions immediately after the end of the colonial period. We use constraints on the executive at independence, measured by PolityIV for the independence of each former colony. A final control is the date of independence. This is useful because constraint on the executive at different dates of independence may mean different things.

## 2.4 The Sample

Table 1 displays the descriptive statistics for our sample in the year 2000. Segregation is an index between 0 and 1. The five least segregated countries are Germany, Sweden, the Netherlands, South Korea and Japan, with figures close to zero. The most segregated country is Zimbabwe with 0.39, followed by Guatemala, Afghanistan, Uganda and Turkey. The mean segregation is 0.10. Regarding conflict, of 91 countries, there is low-, medium- and high-intensity conflict in 20, 16 and 10 of those countries in 2000.[6] For the continuous conflict index, the mean is 0.09, with 40 countries scoring zero and with Mexico, Colombia, Turkey, Sudan and Peru as the more troubled countries in 2000.

*** TABLE 1 ***

An eventual shortcoming of our empirical strategy is that cross-country data for segregation are scant. Matuszeski and Schneider (2006) provide segregation measures for a larger number of countries, but they rely on proprietary data that are not publicly available. Alesina and Zhuravskaya (2012) provide ethnic segregation data for a cross-section of 98 countries; when we consider countries with data on conflict, the sample is further reduced to 91. The obvious concern is whether our results for a subset of countries may be statistically validated on the whole sample or not. An obvious problem emerges if the subsample of countries with data on segregation is very different from the rest of the sample. For instance, information may be more thoughtfully and extensively collected in developed countries.[7] Columns (4) to (6) display the same description for the enlarged sample of countries with data on other variables but not on segregation. We observe that in every dimension, the subset with segregation data is very close in the main statistics to the whole sample. The means do not diverge in more than one-third of their respective standard deviations. While this verification does not avoid other selection problems, at least we are able to observe that our sample is not obviously biased in a particular direction.

[6] Given that the definition of conflict is above a given threshold, high-conflict countries are also intermediate- and low-conflict countries, while intermediate-conflict countries are also low-conflict countries.
[7] Nonetheless, Alesina and Zhuravskaya (2011) report that this is not the case with segregation data.

For our panel data study, we consider data on segregation for the year 2000 and all the fixed variables described in Table 1. The measure of conflict, on the contrary, changes over time.[8] The panel data exhibit the same similarity between the whole sample and the subsample as the available data on segregation. The post-war sample is composed of 90 countries over the period 1960–2005, with a total of 679 observations for the largest sample, with observations in each 5-year sub-period.

## 3. A Cross-National Analysis of Segregation and Conflict

First, we use cross-country regressions to study the correlation between conflict and segregation. Two papers study the effect of segregation in a cross-section of countries. Matuszeski and Schneider (2006) analyze the same correlation, but their measure of segregation is different from ours. Alesina and Zhuravskaya (2012) use their segregation measure to study its effect on governance and other institutional arrangements.

For the construction of the segregation indices, Alesina and Zhuravskaya (2012) draw data closest to the year 2000 whenever its results are available. Accordingly, we use 2000 as our baseline year. For both the dependent variable and the controls we use the data available for that year. While we can use averaged data as well, for instance the number of conflicts by country during the postwar period,[9] we prefer to work only with one year and keep the rest of the longitudinal data for panel regressions.

As an exploration, we construct a simple horse race between the different variables associated with conflict. Together with the segregation index, we include income and democracy in 2000 on the one hand and fractionalization and polarization on the other. We construct a subsample with countries in the last decile of each of those variables, and we count the number of low,

[8] Also changing over time are the GDP, population and democracy.
[9] Matuszeski and Schneider (2006).

intermediate and high conflicts there.[10] We repeat the exercise using the first half of the sample.[11] Table 1 displays the result for each level of conflict.

*** FIGURE 1 ABOUT HERE ***

We observe that 5 out of 20 countries with conflicts are in the last decile of the segregation measure, and 16 out of 20 countries are in the higher half. The figures are similar for the other intensity of conflict, showing that the distribution of segregation among conflict countries is skewed to the right. While this is also the case for all the other variables in the figure, we notice that segregation has a higher likelihood of predicting conflict than all the other variables.

Now we undertake cross-country estimation. Our baseline measure is low-intensity conflict and we also provide results for the continuous measures with several observations censored in zero. Accordingly, we use robust probit and tobit estimators, respectively. Table 2 displays the results.

*** TABLE 2. CROSS-SECTIONAL OLS ***

As shown in Table 2, segregation is significantly and positively correlated with the probability and intensity of conflict. The addition of controls does not affect the sign of the results. The results offer support to the hypothesis that segregation has a positive effect on conflict. More noticeably, columns (1) and (4) indicate that the pure correlation between segregation and conflict is positive. This result is stronger than the one found by both Alesina and Zhuravska (2012) and Matuszeski and Schneider (2006), because in their work the positive (respectively, negative) effect of segregation on conflict (respectively, governance) is obtained only after controlling for fractionalization. In terms of the magnitude of this effect, a change in one standard deviation of segregation increases the probability of conflict by about 10%; countries in the last decile of segregation have a 30% higher probability of conflict.

[10] From 90 countries, we select the 9 countries with higher values for segregation and then we count the number of conflicts within them. We repeat the same for all the variables. In the case of income and democracy, we consider the 9 countries with the lowest figures on these indices.

[11] From 90 countries, now we count conflicts over the 45 countries with higher values for segregation, fractionalization and polarization and lower values for income and democracy, respectively.

The previous results show the strong correlation between segregation and conflict. Advancing to propose a causal interpretation of this correlation, we take advantage of the instrumental variable also constructed by Alesina and Zhuravska (2012). Table 3 replicates the same regressions as Table 2, but using the IV in the first stage.[12]

*** TABLE 3. CROSS-SECTIONAL IV ***

The results exhibited in Table 3 are reassuring about the relation between segregation and conflict. For binary low-intensity conflict the coefficients are higher than in Table 2, and for continuous conflict the figures are similar. We conclude that there is not an evident problem of reverse causality. While other variables are significant in some regressions but not in others, segregation is consistently positive across all estimations. The next section moves onto the panel data analysis of the same relation.

## 4. Panel Analysis

We consider the following dynamic linear specification:

$$c_{it} = \gamma c_{it-1} + \beta X_{it-1} + \delta_i + \delta_t + \varepsilon_{it}$$

where $c_{it}$ is conflict for country $i$ in period $t$, $X_{it}$ are controls, and $\delta_i$ and $\delta_t$ represent country and time fixed effects, respectively. $\varepsilon_{it}$ is a disturbance term, clustered by country. The sample includes all independent countries during the post-war period, with observations taken every fifth year from 1960 to 2000.

Equation (1) is typically used in order to estimate the effect of a time-dependent variable $X_{it}$ on the dependent variable $c_{it}$. In this context, the fixed effect has the purpose of merely controlling for some non-observable heterogeneity. However, in their study of modernization theory, Acemoglu et al. (2009) propose to study the fixed effect as well in terms of some observable

[12] First-stage regressions are not reported. The IV prediction of segregation in the first stage is very strong, with F over 30 if we use a linear 2SLS estimator.

variables that intend to proxy for the long-run determinants included in the fixed effect: "the fixed effects … should be closely linked to the underlying institutional development paths and to the factors affecting what type of path a society has followed." Fixed effects are modeled in terms of fixed-over-time variables, as follows:

$$\delta_i = \alpha S_i + \alpha' z_i + \mu_i$$

with $S_i$ referring to segregation, which we presume to be fixed over time, $z_i$ meaning all other fixed variables and $\mu_i$ an error term. We argue that segregation is a main component of $\delta_i$ in (1), which implies a significant value of $\alpha$ in (2). To give a preliminary illustration of this argument, we run a simple OLS of (1) to obtain an estimation of $\delta_i$ [13] and then we relate this estimated fixed effect to segregation. Figure 2 shows this relation, conditioning by geographical covariates and other ethnic variables.

*** FIGURE 2 ABOUT HERE ***

We observe that segregation is strongly correlated with fixed effects. In fact, segregation has a positive and significant correlation with the fixed effect after controlling for all the other variables.[14]

To proceed with the panel estimations, we introduce (2) into (1). The error term $\mu_i$ implies that the estimation method is a random-effect regression. We use the same type of estimators as in the cross-section, namely a probit for binary conflict and a tobit for continuous conflict censored at zero. Firstly, we add the same controls as in the cross-country analysis. Table 4 displays our baseline regressions for panel data.

*** TABLE 4 ABOUT HERE ***

[13] The OLS estimation avoids the problem of sample reduction due to non-linear estimation, such as probit or logit, in the case of binary conflict. In such a case, all countries where conflict is not changing are simply dropped out of the sample and we are unable to estimate the country-specific effect. While this estimation has a problem of consistency, since the number of estimations increases with the number of countries, we only use it to provide an illustration before undertaking consistent estimations. The same strategy is used by Acemoglu et al. (2009).

[14] These regressions are available upon request. However, they should be interpreted merely as an illustration, since fixed effects in linear models are not estimated consistently.

We observe that segregation is positive and significant in all the regressions. The pure coefficient in column (1) indicates that a change of one standard deviation of segregation increases the probability of conflict by about 5%. Compared with cross-country regression, this coefficient is lower, which is not a surprise because now the effect is conditional on conflict in the previous period. However, the coefficients for regressions with controls (columns 3 and 4) and for the continuous measure of conflict are similar to those in Table 2. Regarding the controls, only Ncont is consistently positive and significant for geography, while both population and income have the expected signs.

A major problem with panel regression such as the one reported in Table 4 is omitted variable bias. Even though we are controlling for a number of variables as a proxy for the fixed effect (2), there are some long-run variables that can be contained in $\mu_i$ and correlated with conflict.
Acemoglu et al. (2001) document that factors affecting the profitability of different institutional structures for European colonizers had a major impact on early institutions and on subsequent political and economic development in former European colonies. Accordingly, these institutional arrangements can bias the random-effect panel regressions (Acemoglu et al., 2009).[15] To avoid this possibility, we undertake the same strategy, namely to condition the post-war estimations for a number of institutional variables. We use the log population density at 1500, years since independence and polity at independence as our proxy for long-run institutions. These control variables provide an additional test of robustness since they are reducing the same in a half.[16] Table 5 displays our results.

*** TABLE 5 ABOUT HERE ***

[15] These authors suggest that the well-known relation between democracy and income can be given by these long-run effects.
[16] We have positive but weaker results using an additional variable: settler mortality. However, in this case, the sample is further reduced.

The results in Table 5 are reassuring about the relation between segregation and conflict. While the institutional variables have positive and significant coefficients on the estimation, the effect of segregation is preserved. This means that even conditional to the institutional structure, segregation has an important correlation with conflict.

## Conclusions

This work represents an effort to cast new light on the relationship between segregation and social conflict. The traditional literature on social conflict has identified as its main drivers the extent of ethnic, linguistic and religious polarization and fractionalization. The reason is that in order to articulate a given social group, economic characteristics are not enough: a poor group will find it very difficult to organize itself due to the lack of resources. Consequently, another aspect is needed for groups' internal organization, such as religion, race or language. With these elements in mind, several investigations have been conducted using distributional measures of these aspects to explain social conflict. As mentioned, the two most important measures used are the fractionalization and polarization indices. Despite their importance, they do not consider the groups' location and how they share the space within a region. If two groups live apart, their preferences can diverge greatly from each other and articulate themselves more easily. The latter can increase the possibility of a conflict. Despite the potential relevance of segregation as a driver of conflict, little has been undertaken to try to understand the relationship between these two phenomena. In this article, we have attempted to add new insights to this analysis; in particular, we have tried to estimate the relationship between segregation and conflict. Using a panel of 91 countries from1950 until 2000, we found robust evidence that this relationship between segregation and social conflict is strong. This result survives even after controlling for polarization and fractionalization.

**References**


Acemoglu, D., Johnson, S. and Robinson, J. (2001) "The colonial origins of comparative development: An empirical investigation," *American Economic Review* 91, 1369–1401.

Acemoglu, D., Johnson, S., Robinson, J. and Yared, P. (2009) "Reevaluating the modernization hypothesis," *Journal of Monetary Economics* 56, 1043–1058.

Alesina, A. and LaFerrara, E. (2000) "Participation in heterogeneous communities," *Quarterly Journal of Economics* 115(3), 847–904.

Alesina, A. and Zhuravskaya, E. (2012) "Segregation and the quality of government in a cross section of countries," *American Economic Review* 101(5).

Alvarez, M., Cheibub, J., Limongi, F. and Przeworski, A. (2000) "Democracy and Development: Political Institutions and Material Well-Being in the World, 1950–1990," Cambridge University Press.

Anas, A. (2002) "Prejudice, exclusion, and compensating transfers: The economic of ethnic segregation," *Journal of Urban Economics* 52, 409–432.

Burnell, J. (1988) "Crime and racial composition in contiguous communities as negative externalities: Prejudiced households' evaluation of crime rate and segregation nearby reduces housing values and tax revenues," *Journal of Economics and Sociology* 47(2), 177–193.

Burton, E. (2003) "Housing for an urban renaissance: Implications for social equity," *Housing Studies* 18(4), 537–562.

Charles, C., Dinwiddie, G. and Massey, D. (2004) "The continuing consequences of segregation: Family stress and college academic performance," *Social Science Quarterly* 85(5), 1353–1373.

Cheshire, P. (2007) “Segregated Neighbourhoods and Mixed Communities: A Critical Analysis,” Joseph Rowntree Foundation.

Clapp, J. and Ross, S. (2004) “Schools and housing markets: An examination of school segregation and performance in Connecticut,” *Economic Journal* 114(499), 425–440.

Collier, P. and Hoeffler, A. (1998) “On economic causes of civil war,” *Oxford Economic Papers* 50, 563–573.

Collier, P. and Hoeffler, A. (2002) “On the incidence of civil war in Africa,” *Journal of Conflict Resolution* 46(1), 13–28.

Collier, P. and Hoeffler, A. (2004) “Greed and grievance in civil war,” *Oxford Economic Papers* 56(4), 563–95.

Collier, P., Hoeffler, A. and Soderbom, M. (2004) “On the duration of civil war,” *Journal of Peace Research* 41, 253–273.

Conejeros, R. and Vargas, M. (2007) “Group effect, productivity and segregation optimality,” *Lecture Notes in Economics and Maathematical Systems* 599, 209–222.

Conejeros, R. and Vargas, M. (2012) “Segregation, exclusion and compensating transfers under a dynamic setting,” *Applied Economics* 44(17), 2203–2215.

Dawkins, J., Shen, Q. and Sanchez, T. (2005) “Race, space, and unemployment duration,” *Journal of Urban Economics* 58, 91–113.

Dosh, P. (2003) “Review essay: Violence, spatial segregation, and the limits of local empowerment in urban Latin America,” *Latin American Politics and Society* 45(4), 129–146.

Doyle, M. and Sambanis, S. (2003) “International peacebuilding: A theoretical and quantitative analysis,” *American Political Science Review* 94(4), 779–801.

Easterly, W. and Levine, R. (2003) “Tropics, germs, and crops: How endowments influence economic development,” *Journal of Monetary Economics* 50(1), 3–39.

Esteban, J. and Ray, D. (1999) “Conflict and distribution,” *Journal of Economic Theory* 87(2), 379–415.

Esteban, J. and Ray, D. (2008) “On the salience of ethnic conflict,” *American Economic Review* 98(5), 2185–202.

Esteban, J. and Ray, D. (2008) “Polarization, fractionalization and conflict,” *Journal of Peace Research* 45, 163–182.

Esteban, J.L.M. and Ray, D. (2012) “Ethnicity and conflict: An empirical study,” *American Economic Review* 102(4).

Esteban, J. and Ray, D. (2011) “Linking conflict to inequality and polarization,” *American Economic Review* 101(4), 1345–74.

Fearon, J. (2003) “Ethnic and cultural diversity by country,” *Journal of Economic Growth* 8(2), 195–222.

Fearon, J. and Laitin, D. (2003) “Ethnicity, insurgency and civil war,” *American Political Science Review* 97, 75–90.

Fearon, J. (2005) “Primary commodity exports and civil war,” *Journal of Conflict Resolution* 49(4), 483–507.

Giddens, A. (1984) “The Constitution of Society: Outline of the Theory of Structuration,” University of California Press.

Glaeser, E., Laibson, E. and Sacerdote, B. (2002), “An economic approach to social capital,” *Economic Journal* 112(483), 437–458.

Gleditsch, N., Wallensteen, P., Eriksson, M., Sollenberg, M. and Strand, H. (2002) “Armed conflict 1946–2001: A new dataset,” *Journal of Peace Research* 39(5), 615–37.

King, T. and Mieszkowski, P. (1973) “Racial discrimination, segregation, and the price of housing,” *Journal of Political Economy* 81, 590–606.

Kirby, A. (1989) “A sense of place, critical studies in mass communication,” *Media and Community* 6(3), 322–326.

Knight, J. (1992) “Institutions and Social Conflict,” Cambridge University Press.

Krugman, P. (1996) “The Self-Organizing Economy,” Blackwell, Oxford.

Landes, D. (1998) “The Wealth and Poverty of Nations,” W.W. Norton and Company, Inc.

LaVeist, T. (2003) “Racial segregation and longevity among African Americans: An individual-level analysis,” *Health Services Research* 38, 1719–1733.

Logan, J. and Messner, S. (1987) “Racial residential segregation and suburban violent crime,” *Social Science Quarterly* 68(3), 510–527.

Maddison, A. (2008) “Historical statistics of the world economy: 1–2008 AD: Dataset,” http://www.ggdc.net/maddison/Maddison.htm (accessed October 1, 2010).

Madden, J. (2001) "Do racial composition and segregation affect economic outcomes in metropolitan areas?" Problem of the Century: Racial Stratification in the United States, Russell Sage Foundation.

Massey, D. (2001) "Segregation and violent crime in urban America," Problem of the Century: Racial Stratification in the United States, Russell Sage Foundation.

Matuszeski, J. and Schneider, F. (2006) "Patterns of ethnic segregation and civil conflict," Mimeo, Harvard University.

Miguel, E., Satyanath, S. and Sergenti, E. (2004) "Economic shock and civil conflict: An instrumental variables approach," *Journal of Political Economy* 112(4), 725–53.

Molina, G., Pérez de Rada, E. and Jiménez, W. (2002) "Social exclusion: Residential segregation in Bolivian cities," *Research Network Working Papers*, Inter American Development Bank.

Montalvo, J. and Reynal-Querol, M. (2005) "Ethnic polarization, potential conflict, and civil war," *American Economic Review* 95(3), 796–816.

Novta, N. and Klasnja, M. (2012) "Segregation, polarization and ethnic conflict," Mimeo, Department of Economics, New York University.

Portes, A. (1998) "Social capital: Its origins and applications in modern sociology," *Annual Review of Sociology* 24, 1–24.

Ross, M. (2012) "The Oil Curse: How Petroleum Wealth Shapes the Development of Nations," Princeton University Press.

Rummel, R.J. (1963) "Dimensions of conflict behavior within and between nations," *General Systems Yearbook of the Society for General Systems Research* 8, 1–50.

Sachs, J. and Warner, A. (1997) “Fundamental sources of long-run growth,” *American Economic Review Papers and Proceedings* 87, 184–188.

Schelling, T. (1971) “Dynamic models of segregation,” *Journal of Mathematical Sociology* 1, 143–86.

Schneider, G. and Wiesehomeier, N. (2006) “Ethnic polarization, potential conflict, and civil wars: Comment,” unpublished.

Skirmuntt, M. (2012) “Ethnic Segregation and Collective Political Action in Urban England,” Department of Government, University of Essex.

Strand, H. (2006) “Onset of armed conflict: A new list for the period 1946–2004, with applications,” unpublished.

Weidmann, N. (2009) “Geography as motivation and opportunity: Group concentration and ethnic conflict,” *Journal of Conflict Resolution* 53, 526–543.

Weidmann, N. (2011) “Violence ‘from above’ or ‘from below’? The role of ethnicity in Bosnia’s civil war,” *Journal of Politics* 73(4), 1–13.

Wilson, F. and Hammer, R. (2001) “Ethnic residential segregation and its consequences,” Urban Inequality: Evidence from Four Cities, Russell Sage Foundation.

Yinger, J. (2001) “Housing discrimination and residential segregation as causes of poverty,” Understanding Poverty, Harvard University Press.

Young, H.P. (1998) “Individual Strategy and Social Structure. An Evolutionary Theory of Institutions,” Princeton, NJ: Princeton University Press.

TABLES AND FIGURES

FIGURE 1

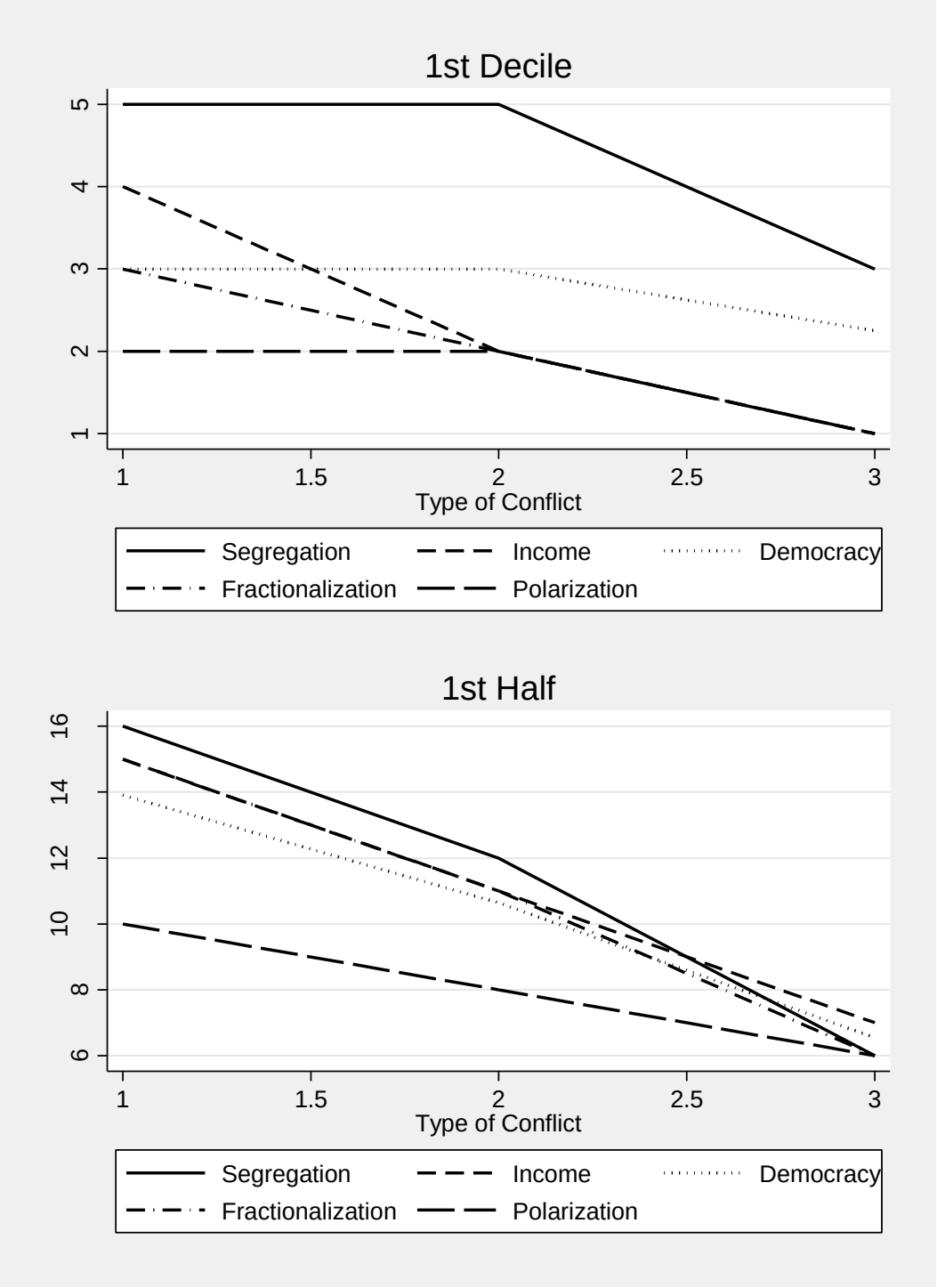

FIGURE 2

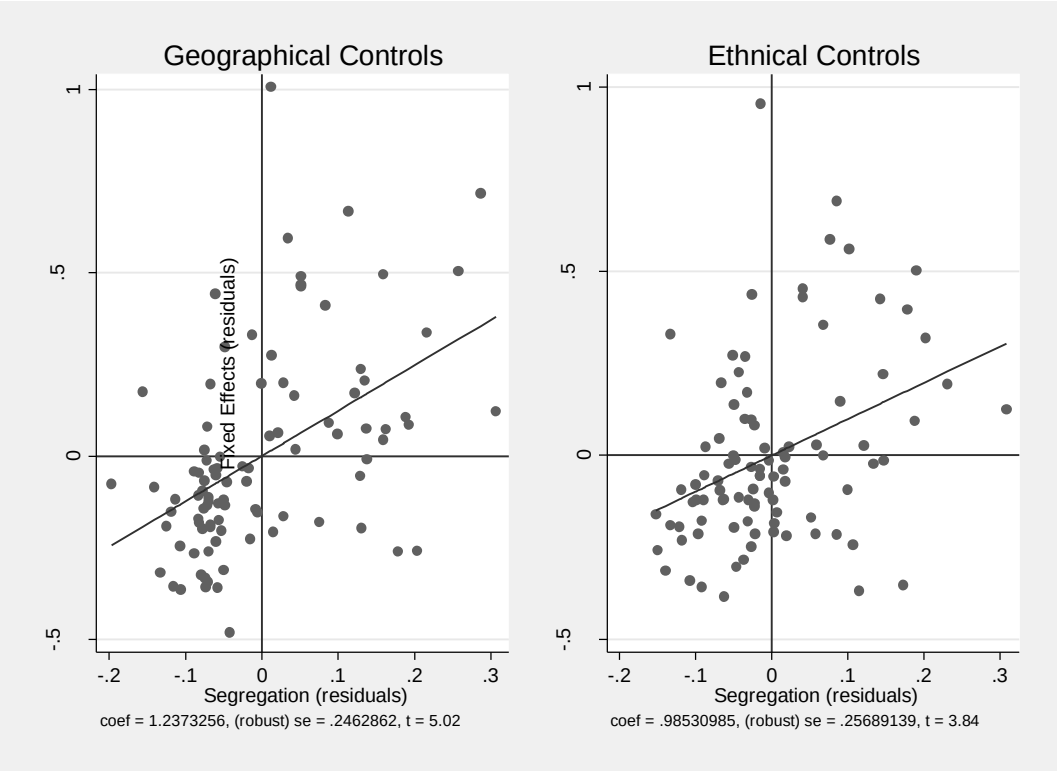
Geographical Controls
Fixed Effects (residuals)
1
.5
0
-.5
-.2
-.1
0
.1
.2
.3
Segregation (residuals)
coef = 1.2373256, (robust) se = .2462862, t = 5.02
Ethnical Controls
1
.5
0
-.5
-.2
-.1
0
.1
.2
.3
Segregation (residuals)
coef = .98530985, (robust) se = .25689139, t = 3.84

TABLE 1

TABLE 1. DESCRIPTIVE STATISTICS

| | Sample | | | Extended Sample | | |
|---|---|---|---|---|---|---|
| | obs. | mean | s.d. | obs. | mean | s.d. |
| Segregation | 91 | 0.10 | 0.11 | - | - | - |
| Conlict Low | 91 | 0.22 | 0.42 | 159 | 0.23 | 0.42 |
| Conflict Med | 91 | 0.18 | 0.38 | 159 | 0.18 | 0.39 |
| Conflict Hight | 91 | 0.11 | 0.31 | 159 | 0.11 | 0.31 |
| Conflict Cont. | 91 | 0.09 | 0.15 | 159 | 0.09 | 0.14 |
| Mountain | 91 | 19.00 | 21.93 | 150 | 16.04 | 20.20 |
| Non Contigous | 91 | 0.19 | 0.39 | 150 | 0.15 | 0.36 |
| Oil/Diamond | 91 | 0.29 | 0.45 | 150 | 0.34 | 0.48 |
| Log Population | 90 | 9.67 | 1.43 | 151 | 9.33 | 1.45 |
| Log Income | 90 | 8.37 | 1.22 | 151 | 8.13 | 1.17 |
| Democracy | 90 | 0.76 | 0.29 | 151 | 0.65 | 0.33 |
| Fractionalization | 90 | 0.39 | 0.25 | 140 | 0.41 | 0.24 |
| Polarization | 90 | 0.04 | 0.05 | 140 | 0.05 | 0.05 |

TABLE 2

TABLE 2. CROSS NATIONAL REGRESSIONS

| Conflict | Conflict is Binary | | | | Conflict is Continuos | | | |
|---|---|---|---|---|---|---|---|---|
| Segregation | 1.164*** | 1.096*** | 0.721* | 0.649 | 0.689*** | 0.605*** | 0.604*** | 0.439** |
| | (0.38) | (0.39) | (0.40) | (0.47) | (0.19) | (0.15) | (0.18) | (0.19) |
| Mountains | | 0.003 | 0.003 | 0.003** | | 0.002*** | 0.002** | 0.001* |
| | | (0.00) | (0.00) | (0.00) | | (0.00) | (0.00) | (0.00) |
| Ncont | | 0.235* | 0.344** | 0.359** | | 0.091** | 0.018 | 0.039 |
| | | (0.13) | (0.16) | (0.17) | | (0.04) | (0.06) | (0.05) |
| Oil/Diamond | | 0.086 | 0.043 | 0.024 | | 0.054 | 0.044 | 0.048 |
| | | (0.10) | (0.09) | (0.08) | | (0.05) | (0.04) | (0.04) |
| Log_Population | | | 0.035 | 0.034 | | | 0.037** | 0.039*** |
| | | | (0.03) | (0.03) | | | (0.02) | (0.01) |
| Log_Income | | | -0.143*** | -0.119*** | | | -0.016 | -0.022 |
| | | | (0.04) | (0.05) | | | (0.02) | (0.02) |
| Democracy | | | 0.234 | 0.239 | | | 0.153* | 0.119* |
| | | | (0.17) | (0.19) | | | (0.08) | (0.07) |
| Fractionalization | | | | 0.233 | | | | -0.021 |
| | | | | (0.18) | | | | (0.08) |
| Polarization | | | | -0.638 | | | | 0.977* |
| | | | | (1.08) | | | | (0.52) |
| Countries | 90 | 90 | 90 | 90 | 90 | 90 | 90 | 90 |

TABLE 3

TABLE 3. CROSS NATIONAL IV REGRESSIONS

| Conflict | Conflict is Binary | | | | Conflict is Continuos | | | |
|---|---|---|---|---|---|---|---|---|
| Segregation | 1.238* | 1.532** | 1.504** | 1.662** | 0.789** | 0.883*** | 0.749*** | 0.613* |
| | (0.712) | (0.784) | (0.802) | (0.902) | (0.321) | (0.307) | (0.270) | (0.343) |
| Mountains | | 0.003 | 0.002 | 0.003 | | 0.002** | 0.002** | 0.001* |
| | | (0.002) | (0.002) | (0.002) | | (0.001) | (0.001) | (0.001) |
| Ncont | | 0.255** | 0.367** | 0.357** | | 0.100*** | 0.017 | 0.034 |
| | | (0.124) | (0.144) | (0.154) | | (0.035) | (0.058) | (0.054) |
| Oil/Diamond | | 0.066 | 0.012 | 0.003 | | 0.044 | 0.041 | 0.046 |
| | | (0.097) | (0.080) | (0.077) | | (0.048) | (0.044) | (0.040) |
| Log_Population | | | 0.033 | 0.028 | | | 0.036** | 0.038*** |
| | | | (0.029) | (0.030) | | | (0.015) | (0.014) |
| Log_Income | | | -0.126 | -0.113 | | | -0.013 | -0.021 |
| | | | (0.044) | (0.044) | | | (0.016) | (0.016) |
| Democracy | | | 0.252 | 0.266 | | | 0.161** | 0.130** |
| | | | (0.173) | (0.182) | | | (0.076) | (0.065) |
| Fractionalization | | | | 0.085 | | | | -0.048 |
| | | | | (0.206) | | | | (0.095) |
| Polarization | | | | -1.014 | | | | 0.897 |
| | | | | (1.217) | | | | (0.571) |
| Countries | 90 | 90 | 90 | 90 | 90 | 90 | 90 | 90 |

TABLE 4

TABLE 4. PANEL REGRESSIONS

| Conflict | Conflict is Binary | | | | Conflict is Continuos | | | |
|---|---|---|---|---|---|---|---|---|
| Lagged Conflict | 0.601*** | 0.580*** | 0.538*** | 0.521*** | 0.642*** | 0.568*** | 0.507*** | 0.491*** |
| | (0.06) | (0.06) | (0.06) | (0.07) | (0.05) | (0.05) | (0.05) | (0.05) |
| Segregation | 0.762*** | 0.791*** | 0.652*** | 0.499*** | 0.199*** | 0.235*** | 0.212*** | 0.151** |
| | (0.16) | (0.16) | (0.17) | (0.17) | (0.05) | (0.05) | (0.06) | (0.06) |
| Mountains | | 0.001* | 0.001 | 0.001 | | 0.001*** | 0.001** | 0.001* |
| | | (0.00) | (0.00) | (0.00) | | 0.00 | 0.00 | 0.00 |
| Ncont | | 0.106* | 0.098 | 0.146* | | 0.074*** | 0.046** | 0.059*** |
| | | (0.06) | (0.07) | (0.08) | | (0.02) | (0.02) | (0.02) |
| Oil/Diamond | | 0.047 | 0.048 | 0.026 | | 0.02 | 0.017 | 0.015 |
| | | (0.04) | (0.04) | (0.04) | | (0.02) | (0.01) | (0.02) |
| Log_Population | | | 0.023 | 0.023* | | | 0.022*** | 0.023*** |
| | | | (0.01) | (0.01) | | | (0.01) | (0.00) |
| Log_Income | | | -0.068*** | -0.061** | | | -0.015* | -0.016* |
| | | | (0.02) | (0.03) | | | (0.01) | (0.01) |
| Democracy | | | 0.114 | 0.105 | | | 0.047** | 0.044* |
| | | | (0.08) | (0.07) | | | (0.02) | (0.02) |
| Fractionalization | | | | 0.158** | | | | 0.016 |
| | | | | (0.08) | | | | (0.03) |
| Polarization | | | | 0.340 | | | | 0.326** |
| | | | | (0.39) | | | | (0.14) |
| Observations | 613 | 613 | 613 | 613 | 534 | 534 | 534 | 534 |

TABLE 5

TABLE 5. PANEL REGRESSIONS CONTROLLING FOR INSTITUTIONS

| Conflict | Conflict is Binary | | | | Conflict is Continuos | | | |
|---|---|---|---|---|---|---|---|---|
| Lagged Conflict | 0.479*** | 0.444*** | 0.429*** | 0.420*** | 0.641*** | 0.576*** | 0.540*** | 0.534*** |
| | (0.06) | (0.05) | (0.05) | (0.06) | (0.05) | (0.07) | (0.07) | (0.07) |
| Segregation | 0.651*** | 0.695*** | 0.734*** | 0.625** | 0.127* | 0.162** | 0.178** | 0.149* |
| | (0.24) | (0.25) | (0.27) | (0.27) | (0.07) | (0.08) | (0.09) | (0.08) |
| Year Independence | -0.043 | -0.051 | -0.032 | -0.002 | -0.057*** | -0.067*** | -0.060** | -0.052* |
| | (0.04) | (0.05) | (0.07) | (0.08) | (0.02) | (0.02) | (0.03) | (0.03) |
| Polity at Independence | 0.289** | 0.245** | 0.194 | 0.171 | 0.116*** | 0.100*** | 0.084** | 0.080** |
| | (0.12) | (0.12) | (0.15) | (0.16) | (0.04) | (0.03) | (0.04) | (0.04) |
| Logpop Density 1500 | 0.081*** | 0.092*** | 0.072*** | 0.068*** | 0.015** | 0.020** | 0.01 | 0.009 |
| | (0.02) | (0.02) | (0.03) | (0.02) | (0.01) | (0.01) | (0.01) | (0.01) |
| Mountains | | 0.001 | 0.001 | 0.001 | | 0 | 0 | 0 |
| | | (0.00) | (0.00) | (0.00) | | 0.00 | 0.00 | 0.00 |
| Ncont | | 0.282* | 0.205 | 0.23 | | 0.072 | 0.044 | 0.042 |
| | | (0.16) | (0.17) | (0.18) | | (0.05) | (0.04) | (0.05) |
| Oil/Diamond | | 0.099* | 0.075 | 0.083 | | 0.045** | 0.032 | 0.037* |
| | | (0.06) | (0.06) | (0.06) | | (0.02) | (0.02) | (0.02) |
| Log_Population | | | 0.04 | 0.041 | | | 0.020*** | 0.021*** |
| | | | (0.03) | (0.03) | | | (0.01) | (0.01) |
| Log_Income | | | -0.045 | -0.052 | | | -0.019 | -0.024* |
| | | | (0.05) | (0.05) | | | (0.02) | (0.01) |
| Democracy | | | 0.134 | 0.115 | | | 0.04 | 0.031 |
| | | | (0.14) | (0.15) | | | (0.03) | (0.04) |
| Fractionalization | | | | -0.006 | | | | -0.023 |
| | | | | (0.17) | | | | (0.04) |
| Polarization | | | | 0.692 | | | | 0.264** |
| | | | | (0.47) | | | | (0.13) |
| Observations | 333 | 333 | 333 | 333 | 291 | 291 | 291 | 291 |