Diversity and violence during conflict migration: The Troubles in Northern Ireland

Claire L. Adida1*, Joseph M. Brown2, Gordon C. McCord3 and Paul McLachlan4

1Political Science, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA, 2Political Science, University of Massachusetts Boston, Boston, USA, 3School of Global Policy and Strategy, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA and 4SimBioSys Lab, Emory University, Atlanta, USA

*Corresponding author. Email: cadida@ucsd.edu

(Received 13 May 2021; revised 11 February 2022; accepted 27 May 2022; first published online 26 September 2022)

Abstract

Diversity’s effect on violence is ambiguous. Some studies find that diverse areas experience more violence; others find the opposite. Yet conflict displaces and intimidates people, creating measurement challenges. We propose a novel indicator of diversity that circumvents these problems: the location of physical structures at disaggregated geographical levels. We introduce this solution in the context of the Troubles in Northern Ireland. Our data reveal a curvilinear relationship between diversity and conflict-related deaths, with the steepest increase at low diversity, driven by an increase in violence when our proxy for the Catholic proportion of the population rises from 0 to 20 percent. These patterns are consistent with a theory of group threat through exposure.

Keyword: Measurement; conflict; ethnicity; migration

Conflict between Catholics and Protestants in Northern Ireland dates back to the settlement of the Plantation of Ulster in 1609. In the late 1960s, however, the violence escalated to unprecedented levels, leaving approximately 3500 people dead, of whom half were civilians and 7 percent were children. In addition to intercommunal rioting, the conflict involved three armed blocs organized for violence. Largely Catholic Irish Republican paramilitaries lashed out at the Protestant-dominated government and police, as well as the British military forces sent to root out Republicans and pacify Catholic neighborhoods (Bell, 1997). Protestants formed their own paramilitary groups to resist, in their words, “the forces of Romanism” operating under the guise of the Irish independence movement (UVF, 1971). This 30-year episode, commonly known as the Troubles, has been described as “one of the 20th century’s most violent and enduring sectarian conflicts” (Gladstone and Robins, 2021).

Expert analyses of violence in Belfast, where the majority of deaths from the Troubles occurred, document two distinct phenomena: during this period of instability, violence occurred both where Protestants and Catholics interfaced and within their enclaves. Indeed, before the explosion of the Troubles, Catholics and Protestants in Belfast had organized their lives through a dual geographic arrangement: the sanctuary/interface complex. This spatial organization combined a front line space, where ideological and ethnic contestation occurred, with an ethnically homogeneous enclave or sanctuary space where everyday life took place (Feldman, 1991).

However, in the evolution of the Troubles throughout the early 1970s, these enclaves turned deadly; paramilitaries converted them into operational bases and once-safe spaces became loci...
of violence (Feldman, 1991). In other words, the enclaves offered both a place of refuge and a safe haven from which to organize attacks (Boal, 1972; Doherty and Poole, 1997). This intuition was later tested in a systematic and quantitative manner by Cunningham and Gregory (2014), who found that it was the homogeneous sanctuary spaces that were most deadly during the conflict. Mueller et al. (2019) confirm that violence occurred both in the homogeneous sanctuaries and in the interfaces where the two ethnic groups met.

This duality encapsulates a debate animating empirical studies of ethnic conflict intensity. Some studies find that diversity (defined in this particular case as the collocation of two ethnic groups with roughly equal population shares) predicts greater violence (Posen, 1993; Toft, 2003; Weidmann, 2011; Costalli and Moro, 2012; Balcells et al., 2016). Other studies find the exact opposite, that diverse areas are less violent (Mesev et al., 2009; Fjelde and Hultman, 2014; Kasara, 2016; Hägerdal, 2019; Mueller et al., 2019).

It is noteworthy that disaggregated subnational studies of ethnic civil war tend to share a common problem: they rely on demographic measures collected after conflict onset (Mesev et al., 2009; Cunningham and Gregory, 2014; Hägerdal, 2019; Mueller et al., 2019). These measures are confounded by large population movements that occur in a conflict’s early days, limiting our ability to draw inferences about the role that demographic patterns played during the violence. This is particularly true in the context of the Troubles, which broke out in 1969 with “one of the largest movements of civilian populations in postwar Europe” (Feldman, 1991). These problems are compounded by other difficulties that lead to the systematic undercounting of minority groups in ethnic conflict settings (Kopstein and Wittenberg, 2011). Persecuted ethnic groups fearful for their safety may resist the census and violent actors may target census-related infrastructure. Both phenomena occurred in our study area of Belfast (Fermanagh Herald, 1971; Gallagher, 1971, 15; O’Connor, 1971; Sunday Independent, 1971, 1). The 1971 Northern Ireland census used in existing studies (Cunningham and Gregory, 2014; Mueller et al., 2019) likely suffers from widespread measurement error, the typical effect of which is to bias the estimated relationship between independent and dependent variables—in this case, diversity and violence—toward zero (Wooldridge, 2019). Even studies using preconflict data risk a similar bias, because at-risk minority groups may be reluctant to reveal their ethnicity to census enumerators, particularly in demographically mixed areas where tension is high (Silva, 1999, 2007). Studies arguing that diversity intensifies conflict may thus be understating the strength of their own results (Costalli and Moro, 2012).

We propose a solution to this problem with a novel measure of demographic patterns that is not confounded by violence, census resistance, or population movements in the early days of a conflict: the geographic location of churches. We present an original dataset of 264 churches—225 Protestant and 39 Catholic—in the city of Belfast when the conflict broke out in 1969. The churches in our dataset were constructed in the 17th–20th centuries, and we excluded any church built after 1969. (No churches were destroyed by the violence.) The Industrial Revolution, particularly the growth of Belfast’s linen and shipbuilding industries, initiated a massive expansion of the city’s urban population, from 53,000 in 1831 to over 400,000 a century later (Northern Ireland Statistics and Research Agency, 2005). This population growth triggered a construction boom in residential housing and religious facilities, as employers commissioned affordable housing near their mills and religious orders sought to provide sufficient pews within easy walking distance of Belfasters’ homes (Bardon, 1982; Irvine, 1991; McGee, 2013). Thus, the

---

1Our study focuses on a simple case with two ethnic groups. We discuss extensions to cases with more numerous ethnic groups in the concluding section.

2Following Hancock (1998), we treat Catholic and Protestant groups in Northern Ireland as ethnic identities: “The Northern Ireland Troubles fit the definition of an ethnic conflict between Catholic nationalists and Protestant unionists, each side with its own distinct ancestry, cultural traditions, and sense of identity.”
locations of the churches in our dataset provide a snapshot of Protestant and Catholic settlement patterns, which long preceded the Troubles.

Unlike the 1971 census data used in other studies, our measure of Protestant and Catholic settlement is not confounded by census resistance or the population dislocations that occurred in 1969 and 1970 as the conflict flared up. We join our new church diversity measure with existing data on conflict deaths during the Troubles (CAIN, 2014). These data allow us to identify empirical patterns that do not suffer from measurement error and bias due to conflict-driven changes in population.

Our findings are twofold. First, we show that church diversity and conflict deaths have a curvilinear relationship, with the steepest positive slope in the bottom third of church diversity. In other words, the enclaves were not especially violent, in and of themselves; rather, a marginal increase in diversity near the most homogeneous areas had the strongest effect on violence. Second, our data reveal that the main driver of this pattern is the incremental increase in the Catholic proportion of church density, particularly at low levels. The greatest marginal increase in violent conflict-related deaths occurs as the Catholic proportion of church density increases from zero toward 20 percent. There is no similar effect when increasing the Protestant share of church density from very low levels.

Together, we argue, these empirical patterns are consistent with a story of violence motivated by the threat of outgroup exposure. A rich literature in the social sciences demonstrates that exposure to an outgroup, whether it be an ethnic other (Olzak, 1990; Olzak et al., 1996; Enos, 2014; Adida et al., 2015), a refugee (Dinas et al., 2019), or an immigrant (Newman et al., 2012), increases exclusionary attitudes and behaviors toward that outgroup. This is precisely what happened in the lead-up to The Troubles. When Catholics comprised only a small minority of Belfast’s population, interethnic violence was rare—in fact, Protestants contributed much of the funding to build Belfast’s first and second Catholic churches, in 1784 and 1815 (Irvine, 1991). Interethnic relations soured only as the Catholic population grew from this small, preindustrial baseline. As Belfast industrialized and poor Catholics moved to the city for work, the increase in the Catholic population and the perceived presence of a “reservoir” of additional rural Catholics, ready to “invade” Protestants’ neighborhoods and take their jobs, created an environment of perceived scarcity and competition (Boal, 1982; O’Hearn, 1983). The threat to Protestant hegemony was particularly acute in diversifying areas, and this dynamic is reflected in our Troubles-era data: violence is relatively low when all the nearby churches are Protestant, but even a slight increase in the density of Catholic churches predicts a dramatic increase in the number of conflict deaths.

Our paper makes three distinct contributions. First, it builds on the rich literature on the empirical study of conflict intensity (Kalyvas, 2006; Balcells, 2010). Our curvilinear result reconciles claims that ethnic diversity exacerbates violence (Posen, 1993; Costalli and Moro, 2012; Balcells et al., 2016; Schutte, 2017) with competing claims that diversity attenuates conflict (Mesev et al., 2009; Kasara, 2016; Hägerdal, 2019; Mueller et al., 2019). Second, our paper contributes to the nuanced discussion of the deleterious effects of outgroup exposure in contrast to the potentially prejudice-reducing effects of outgroup contact. Because Belfast’s ethnic populations tend to self-segregate socially (Boal, n.d.), violence-reducing social contact and social capital mechanisms do not operate effectively (Allport, 1954; Putnam, 1994; Varshney, 2001; Glaeser, 2005; Jha, 2013). In diverse but socially segregated neighborhoods, exposure to the ethnic other intensifies fears, hatreds, and the resulting conflict violence. Third, our paper proposes a new and methodologically significant technique for measuring ethnic settlement patterns in conflict zones. It is extremely challenging to estimate neighborhood ethnic demographics in civil war contexts because at-risk minorities may be undercounted (Kopstein and Wittenberg, 2011) and violence tends to trigger migration away from diverse areas (Boal, 1982; Irvine, 1991). Both of these problems bias measures of diversity toward zero. Our method allows researchers to recover the underlying population patterns as they existed prior to conflict, enabling unbiased estimates of diversity and its effects on conflict dynamics.
1. Do diverse neighborhoods exacerbate or alleviate conflict intensity?

Ethnic settlement patterns, such as the diversity or homogeneity of a given area, can influence the intensity of violence once civil war breaks out (Weidmann, 2011; Klašnja and Novta, 2014; Hägerdal, 2019). Diversity’s influence may operate through a number of different mechanisms. The historic experience of collocation with an equally large ethnic group generates competitive dynamics and “homelands” that may overlap on a given tract (Toft, 2003; Weidmann, 2011). Population parity gives ethnic armed groups on both sides an expectation that fighting could bring victory and control of the land (Tilly, 1978; Costalli and Moro, 2012). Diversity may also generate security dilemma dynamics between the collocated ethnic groups; mutual vulnerability and suspicion increase the incentive to strike first and continue fighting until the other group is killed or expelled (Posen, 1993).

Other studies argue that diversity decreases conflict intensity via several distinct mechanisms. First, ethnic concentration, as one finds in segregated stronghold areas, facilitates violent collective action (Weidmann et al., 2010). The capacity for violence dissipates as one moves away from stronghold areas, so diverse zones should see less violence. Two previous studies find support for this argument in our research area of Belfast (Mesev et al., 2009; Mueller et al., 2019). Second, ethnic armed groups may prefer to spend their resources targeting one another’s strongholds in symbolic tit-for-tat killings (Fjelde and Hultman, 2014). Third, diversity may enable better information gathering by armed actors (because coethnics are available to serve as informants among the ethnic outgroup) leading to fewer indiscriminate attacks in diverse areas (Hägerdal, 2019).

An entirely different literature rooted in social psychology offers equally ambiguous predictions. On one hand, social contact and social capital theories give us reason to expect less violence in diversely settled areas. A history of interaction may promote intergroup trust and understanding rather than enmity (Allport, 1954). Such intergroup contact can reduce prejudice, decreasing individual incentives for violence (Pettigrew and Tropp, 2008; Hewstone and Swart, 2011; Lemmer and Wagner, 2015; Lowe, Forthcoming; Mousa, 2020). When ethnic groups interact in shared social organizations and economic activities cutting across ethnic lines, “bridging social capital” enables elites to manage tensions and resolve disputes (Putnam, 2000; Varshney, 2001; Glaeser, 2005; Jha, 2013). In Kenya, for example, ethnically diverse areas experienced less violence in the country’s 2008/9 post-election crisis (Kasara, 2016).

At the same time, exposure to outgroups can increase prejudice and exclusion (Rae et al., 2015), particularly among individuals with pre-existing biases (Schieferdecker and Wessler, 2017) or in contexts of socio-economic competition (Olzak, 1990). In a lab-in-the-field experiment examining the integration experience of Muslim immigrants in France, Adida et al. (2015) find that French participants gave less money to Muslim immigrant game partners (relative to Christian immigrant game partners) when they were exogenously placed in a room with more Muslim immigrants around them. In the aftermath of the 2015 Syrian refugee crises, Dinas et al. (2019) show that Aegean islands that randomly received more Syrian refugees saw greater increases in vote shares for the far-right party than did Aegean islands with fewer Syrian refugees. And across the United States’s 50 largest metropolitan areas, Olzak et al. (1996) show that racial contact and competition are correlated with higher levels of ethnic and racial unrest. An implication of this literature is that ethnically diverse areas should see more violence because of this group-threat via exposure effect.

2. A new measure

The disagreement over diversity’s relationship with violence is exacerbated by a problem of measurement error. Conflict displaces people, making it difficult to capture an accurate snapshot of

---

3 Our analysis pertains to ethnic civil war intensity, specifically whether ethnically diverse areas experience more deaths. It also deals with “irregular” or “guerrilla” civil wars, as opposed to “conventional” civil wars with distinct front lines (Balcells and Kalyvas, 2014).
demographic settlement patterns that have not already been shaped by violence. Seeking safety in numbers, residents flee contested areas and relocate to areas controlled by their ethnic group. In Belfast, the August 1969 onset of violence was responsible for at least eight deaths, the destruction of 150 Catholic homes, and the dislocation of 2069 families, mostly from diverse areas (Griffiths, 1971; Coogan, 1993; CAIN, 2014). Within weeks, 75 percent of Belfast residents retrenched to areas that were 90 percent Protestant or 90 percent Catholic (Boal, 1982; Irvine, 1991). Census data collected after conflict onset understate the degree and spatial extent of pre-conflict diversity, biasing the estimated relationship between diversity and violence toward zero.

A related problem involves measurement error in data collected during a conflict—or even prior to it. Many Belfast residents feared that their 1971 census answers would be shared with police. The fear was disproportionately high among Catholics, as priests and Catholic school teachers encouraged their coethnics not to fill out the forms (O’Connor, 1971). Compounding matters, Irish Republican militants encouraged Catholic census resistance and set fire to the Ministry of Finance where completed forms were kept (Fermanagh Herald, 1971; Gallagher, 1971, 15; Sunday Independent, 1971, 1). The resulting measurement error is likely to understate the amount of diversity in the most violent zones, further attenuating the estimated relationship between diversity and conflict intensity.

We propose a new measure of ethnic neighborhood composition that eschews these challenges by relying on the spatial distribution of Catholic and Protestant churches built before 1969. Churches are a useful proxy for population patterns given Belfast’s history. A sixfold population increase during the Industrial Revolution led religious denominations to commission dozens of new churches to keep up with demand for their services.4 With no mass transit system, it was essential to build these structures within walking distance of parishioners’ homes (McGee, 2013).5 Once built, the churches and related amenities such as schools helped to stabilize ethnic settlement patterns, including “residential mixing” (i.e. diversity), which “displayed considerable locational stability” over time (Boal, 1982, 268). Thus, our map of pre-Troubles churches offers a snapshot of where Protestants and Catholics lived before the explosion of violence drove many of them from their homes.

We rely on multiple sources to find the locations of these churches. We first compiled a list of churches that currently hold religious services from the websites of Ireland’s five main Christian denominations: Catholics (39 churches within 3 km of Belfast), Baptists (25), Presbyterians (83), Methodists (42), and the Church of Ireland (75).6 We then consulted historical sources and emailed the parish offices of all denominations to identify derelict or demolished churches that were open before the onset of the Troubles. We confirmed, using parish histories and other secondary sources, that all the churches in our database predate August 1969. We geolocated each church using street addresses from church websites and historic street directories. Using Google Maps’ satellite imagery and public API, we confirmed each structure’s latitude and longitude.7 Figure 1 shows the locations of these churches. Our data on the location of violent conflict-related deaths come from the Conflict Archive on the Internet (CAIN) project at the University of Ulster (CAIN, 2014). Researchers geolocated the 3529 civilian and combatant deaths during the Troubles to specific street addresses, making the data available as a Google Maps layer, which we scraped from the public XML file.8 Figure 2 plots

---

4 Belfast’s population grew from 53,000 in 1831 to over 400,000 a century later (NISRA 2010).
5 This pattern, building new churches in areas of the corresponding religious denomination’s growth, continued into the 20th century. For instance, the website of Saint Agnes’s church, in the West Belfast neighborhood of Andersonstown, explains that “the sudden explosion in house building after the [Second World] war... made the erection of a new church an urgent necessity.” See http://www.stagnesbelfast.com/?page_id=50.
6 These denominations represent 95 percent of churchgoers in Northern Ireland, according to the 2011 census.
7 Our replication data notes list the sources consulted.
Due to a lack of direct witnesses to some of the killings, 224 of the Belfast deaths in the CAIN data are geolocated to addresses where the victims’ bodies were found. To guard against potential bias that would be caused if some bodies were moved prior to their discovery, Appendix Table A5 excludes these 224 deaths. Our results are unchanged. CAIN is not the only source of geospatial data from the Troubles. The Northern Ireland Research Initiative (NIRI) estimates the locations of violent events using machine learning (Loyle et al., 2014). However, the NIRI dataset contains significant errors—for instance, assigning
a density map of the 3379 conflict-related fatalities within the Belfast city limits, surrounding each death in a 500 m bandwidth kernel. The density of deaths is particularly high in West Belfast, a thickly-settled area where working class Protestants and Catholics lived in close proximity (Boal, n.d.). The density of deaths is also high in the city’s commercial center, frequented by Protestants and Catholics alike.

We use the location of churches predating the 1969 onset of the Troubles to construct a pre-conflict measure of ethnic diversity. We estimate kernel densities stretching out from each church, generating 25 m grids of each denomination’s church density per square kilometer in Belfast. For robustness, we construct kernels stretching 1.5, 2, and 3 km from the church and we report results at varying kernel bandwidths. We construct the measure for the ethnic diversity of the population by creating a variable for “church diversity” in each grid cell. Our measure follows the form of Montalvo and Reynal-Querol (2010)’s RQ:

\[
\text{Church Diversity}_i = 1 - \left[ \left( \frac{0.5 - \pi_{\text{Cath}}}{0.5} \right)^2 \pi_{\text{Cath}} + \left( \frac{0.5 - \pi_{\text{Prot}}}{0.5} \right)^2 \pi_{\text{Prot}} \right]_i
\]

where \( \pi_{\text{Cath}} \) denotes the Catholic church density in that location divided by the sum of Catholic and Protestant church densities, while \( \pi_{\text{Prot}} \) has an analogous definition for Protestant church density. This church diversity measure is closely related to the Herfindahl index, which takes the form \( 1 - \frac{RQ}{2} \) for a two-group setting.\(^9\) Crucially, our measure represents church diversity, rather than church density. This is important because a church diversity measurement is not confounded by population density.

Based on our new measure, cells with a value close to 1 have a similar density of Protestant and Catholic churches, indicating a highly diverse community. Cells with a value near zero have significantly more Catholic or more Protestant churches, indicating low diversity. Cells with neither Catholic nor Protestant churches within a given kernel bandwidth have no value for the measure and are not included in regressions at that kernel bandwidth. Figure 3 maps the church diversity measure for the 1.5 km bandwidth. (The number of churches belonging to each denomination is shown in Table 1.) Allowing for differences in population density, we note a visual correspondence between the presence of darker, more diverse zones and the violent dark spots shown in Figure 2. The many overlapping Catholic and Protestant church kernels of West Belfast, in particular, correspond to a high density of deaths in an area that was both thickly and diversely settled. However, it is also important to note that an area such as West Belfast may contain a complex patchwork of mixed Protestant/Catholic blocks, segregated blocks, and contested streets where majority Catholic and majority Protestant blocks abut (Boal, n.d.). Our ability to measure diversity at a highly disaggregated level allows us to determine which settlement patterns predict more violence: segregated enclaves, fully mixed neighborhoods, or neighborhoods where one ethnic group dominates but its hegemony is threatened by the presence of a minority of households from the other ethnic group.

\(^9\)Our findings are robust to using Ethnolinguistic Fractionalization instead—see Appendix Table A7 and Figure A3. In the conclusion, we discuss potential extensions for more than two groups.

over 80 deaths to a single point in central Belfast and over 30 to a field west of Belfast. These locations are unsupported by qualitative accounts of the conflict (McKittrick et al., 2001).
3. Estimation and results

3.1 Summary statistics

Table 2 presents summary statistics on the variables used in the analysis. There are 8258 locations within Belfast included in the 1971 census, from which we source the household asset variables.\(^{10}\) The data on victims have been mapped to those locations. Only 5451 locations have data on assets. Distance to nearest church is calculated with GIS, and together with the church diversity Figure 3. Diversity of churches constructed from Catholic and Protestant kernel densities in Belfast. Darker areas are more mixed in church density.

Table 1. Churches by denomination

<table>
<thead>
<tr>
<th>Denomination</th>
<th>Churches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catholic</td>
<td>41</td>
</tr>
<tr>
<td>Church of Ireland</td>
<td>76</td>
</tr>
<tr>
<td>Baptist</td>
<td>25</td>
</tr>
<tr>
<td>Methodist</td>
<td>42</td>
</tr>
<tr>
<td>Presbyterian</td>
<td>83</td>
</tr>
<tr>
<td><strong>Total Protestant</strong></td>
<td><strong>226</strong></td>
</tr>
</tbody>
</table>

\(^{10}\)Because they are drawn from the 1971 census, these control variables are prone to some of the same measurement problems we describe earlier in this paper. As a result: (a) we run all our regressions with and without these variables; (b) we acknowledge here that we are introducing possible endogeneity on the right-hand side, potentially biasing our coefficient of interest; and (c) we note that conflict migration and intimidation should not have an effect on the description of household assets the way it may affect respondent identification.
variable is aggregated to 100 m grids that match the 1971 census grids. These 100 m grids are our unit of observation.

### 3.2 Church diversity and violence

We begin by exploring the association between church diversity and the location of conflict deaths. Figure 4 shows the association in the raw data, at various bandwidths for the construction of the church diversity measure. For all three bandwidths, the number of deaths per grid cell is lowest in locations where church diversity is near zero (complete homogeneity), and rises significantly starting at a church diversity of approximately 0.2. Deaths peak when church diversity is in the range of 0.4–0.6, then decrease as diversity approaches 1, although violence levels always remain above those at zero church diversity (i.e., homogeneous neighborhoods).

Our dependent variable is a death count, so we employ a negative binomial specification. We specify a quadratic of the church diversity variable and control for the log of distance to the nearest Catholic church and Protestant church to measure separately the association of deaths to church diversity and church proximity. We control for the log of population density and socioeconomic status (SES) in each grid square, given prior research showing that working-class neighborhoods were particularly conflict-prone (Feldman, 1991). Our SES control is a composite, incorporating the first principal component of four census variables: the proportion of households in the grid cell without hot water, the proportion without a bath in the home, the proportion without an indoor shower, and the proportion without an inside toilet. (Combining these measures based on their first principal component is essential to avoid multicollinearity.) To adjust inference for spatial autocorrelation, we cluster standard errors using 1993 electoral ward boundaries (the earliest available as a shapefile). This clustering has no effect on our coefficients, but it offers a more conservative assessment of the statistical significance of our results.

Table 3 shows the results of these negative binomial estimations, with our diversity variable calculated at several bandwidths. (Histograms of diversity for all bandwidths are shown in Appendix Figure A1.) Column 1 estimates the association between the count of victims and the linear and quadratic terms in church diversity. Both diversity terms are calculated at a 1.5 km bandwidth. The 3.036 coefficient on the linear diversity term is significant at \( p < 0.01 \) and the \(-2.756 \) coefficient on the quadratic term is significant at \( p < 0.05 \). Together, these results suggest that cells with a church diversity of 1 have around 32 percent more deaths than those with a church diversity of zero. A one standard deviation increase in diversity (0.32 at the 1.5 km bandwidth) is associated with a doubling (a 99 percent increase) of deaths in a grid cell. The

### Table 2. Summary statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>Victims in grid cell</td>
<td>8258</td>
<td>0.196</td>
<td>0.904</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Church diversity (1.5 km bw)</td>
<td>8196</td>
<td>0.391</td>
<td>0.320</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Church diversity (2 km bw)</td>
<td>8255</td>
<td>0.392</td>
<td>0.299</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Church diversity (3 km bw)</td>
<td>8258</td>
<td>0.472</td>
<td>0.228</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Distance (m) to nearest Protestant church</td>
<td>8258</td>
<td>463.0</td>
<td>325.5</td>
<td>10</td>
<td>2075</td>
</tr>
<tr>
<td>Distance (m) to nearest Catholic church</td>
<td>8258</td>
<td>1111</td>
<td>877.2</td>
<td>11</td>
<td>5846</td>
</tr>
<tr>
<td>Population</td>
<td>8258</td>
<td>71.19</td>
<td>76.66</td>
<td>1</td>
<td>1438</td>
</tr>
<tr>
<td>Population without hot water</td>
<td>5451</td>
<td>0.0851</td>
<td>0.192</td>
<td>0</td>
<td>0.98</td>
</tr>
<tr>
<td>Population without bath or shower in home</td>
<td>5451</td>
<td>0.137</td>
<td>0.288</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Population without inside toilet</td>
<td>5451</td>
<td>0.142</td>
<td>0.287</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

11Likelihood ratio tests show that the negative binomial specification is preferred over the alternative Poisson specification, shown in Appendix Table A1 and Figure A2. The results of the two specifications are substantively similar.


13We calculate the log increase in deaths by taking the exponent of the regression coefficients: \( e^{3.036-2.756} \).
association between violence and church diversity is independent of proximity to the nearest Catholic or Protestant church.

Columns 2 and 3 show the same church diversity variables with larger kernel bandwidths of 2 and 3 km. Consistent with column 1, the coefficients on the linear diversity term are positive and significant at the \( p < 0.01 \) level, while the coefficients on the quadratic term are negative and significant at the \( p < 0.05 \) level.

![Figure 4](https://doi.org/10.1017/psrm.2022.36)

**Figure 4.** Local polynomial fit of number of victims in grid cell on church diversity measure. Dotted lines indicate 95 percent confidence intervals.

**Table 3.** Victims and church diversity

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1) 1.5 km</th>
<th>(2) 2 km</th>
<th>(3) 3 km</th>
<th>(4) 1.5 km</th>
<th>(5) 2 km</th>
<th>(6) 3 km</th>
</tr>
</thead>
<tbody>
<tr>
<td>Church diversity</td>
<td>3.036***</td>
<td>5.579***</td>
<td>4.619***</td>
<td>2.204*</td>
<td>4.099***</td>
<td>4.966***</td>
</tr>
<tr>
<td></td>
<td>(1.104)</td>
<td>(1.084)</td>
<td>(1.699)</td>
<td>(1.136)</td>
<td>(1.181)</td>
<td>(1.709)</td>
</tr>
<tr>
<td>Church diversity squared</td>
<td>−2.756**</td>
<td>−4.418***</td>
<td>−3.020**</td>
<td>−1.432</td>
<td>−2.419**</td>
<td>−2.689*</td>
</tr>
<tr>
<td></td>
<td>(1.148)</td>
<td>(1.215)</td>
<td>(1.413)</td>
<td>(1.083)</td>
<td>(1.229)</td>
<td>(1.390)</td>
</tr>
<tr>
<td>Observations</td>
<td>8,196</td>
<td>8,255</td>
<td>8,258</td>
<td>5,451</td>
<td>5,451</td>
<td>5,451</td>
</tr>
<tr>
<td>SES control</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Notes: Results from negative binomial estimations, with number of victims as the dependent variable, church diversity constructed at 1.5, 2, or 3 km bandwidths. Regressions (4)–(6) add a control for socioeconomic status using the first principal component of variables for proportion of households without hot water, without a bath or shower, without a toilet, and without an inside toilet. All regressions control for log distance to nearest Catholic church, log distance to nearest Protestant church, and log population. Coefficients for these distance, population, and SES control, as well as for the constant term, are not shown. Robust standard errors, in parentheses, are clustered by 116 wards using 1993 boundaries, and significance levels are denoted at conventional levels ***\( p < 0.01 \), **\( p < 0.05 \), *\( p < 0.1 \).

The association between violence and church diversity is independent of proximity to the nearest Catholic or Protestant church.

Columns 2 and 3 show the same church diversity variables with larger kernel bandwidths of 2 and 3 km. Consistent with column 1, the coefficients on the linear diversity term are positive and significant at the \( p < 0.01 \) level, while the coefficients on the quadratic term are negative and significant at the \( p < 0.05 \) level.
significant at the p < 0.05 level or better. Cells with a church diversity of 1 can be expected to have 3–5 times more deaths (based on columns 2 and 3, respectively) compared to cells with zero diversity. A one standard deviation increase in diversity (0.299 at the 2 km bandwidth and 0.228 at the 3 km bandwidth) is associated with 3.6 or 2.5 times more deaths, respectively.

Columns 4, 5, and 6 replicate the analyses shown in the first three columns, with the addition of our composite SES control. These specifications contain roughly 34 percent fewer observations because 2807 census grid cells lack any data on the four original SES variables. However, the results are broadly similar to those shown in columns 1–3. The sign on the linear diversity term is positive and substantively consistent across kernel bandwidths, while the coefficients on the quadratic diversity terms are consistently negative and substantively consistent. There is a loss of statistical significance in the 1.5 km bandwidth model, due to the considerable reduction in observations and the likely endogeneity on the right-hand side (since the SES variables may have been impacted by the violence). Nonetheless, the coefficients on linear and quadratic terms are both statistically significant at the 2 km bandwidth, and the coefficient on the linear term is significant at the 3 km bandwidth. A one standard deviation increase in diversity is associated with a predicted increase of between 75 and 174 percent of the deaths per cell, depending on the kernel bandwidth. Increasing diversity from zero to one predicts a 116 percent increase in deaths per cell at the 1.5 km, and almost ten times more deaths per cell at the 3 km bandwidth.

For comparison with our main results, Appendix Table A2 shows the results of negative binomial regressions using the flawed 1971 census data rather than our church diversity measure. The result is a null effect for both linear and quadratic measures of diversity in a given census grid cell. The coefficients on both variables fail to reach statistical significance, whether or not SES controls are used. Moreover, the substantive effects of diversity are small and inconsistent—either increasing or decreasing deaths, depending on the inclusion of SES controls. These null findings are precisely what one expects when using census data collected during a conflict. As discussed earlier, respondent religion was measured with error and there were broad problems with the census data, likely biasing the estimated relationship between diversity and violence toward zero. In contrast, the clear and consistent results using churches as proxies show the promise of our novel approach to measuring ethnic diversity independent of conflict migration.

We now explore the marginal effects of our negative binomial estimations using church diversity. Figure 5 plots the effects of church diversity on violence for each bandwidth, as given by the negative binomial regressions with SES controls (Table 3, columns 4–6). With the exception of the smallest bandwidth—where there is less variation—the greatest effect on violence of a marginal increase in church diversity occurs when church diversity is in the 0.3–0.5 range. This is consistent with the trend observed in the raw data in Figure 4, where the steepest part of the curve occurs at diversity levels of roughly 0.4.

In Figure 6, we further investigate what drives this curvilinear pattern by separating out the effect of Catholic versus Protestant diversification. Indeed, an increase in church diversity from 0.3 to 0.4 may be driven either because the share of Catholic churches increased or because the share of Protestant churches increased, or both. Most interpretations of diversity indices either are silent on the particular underlying configuration, or assume symmetry. Yet a theory of group threat would predict a rise in violence in the first scenario (where the population share of the stigmatized Catholic minority increases), not the second (where the population share of the Protestant majority increases). Figure 6 investigates these patterns, showing the predicted level of violence for each 10 percent bin of the Catholic (top three figures) or Protestant (bottom three figures) proportion of church diversity. (The subfigures representing the Catholic proportion are mirror images of the corresponding Protestant subfigures.)

At the 3 km bandwidth, there are no locations over 80 percent Catholic by the church diversity measure, so no probabilities are calculated for the 80–90 or 90–100 percent Catholic (0–10 or 10–20 percent Protestant) bins. The same is true at the 2 km bandwidth for the 0–10 percent Catholic (90–100 percent Protestant) bins.
A marginal increase in predicted deaths occurs as Catholic church density increases from 0–10 to 10–20 percent. The same trend is not observed when Protestants’ share of church diversity increases along that range. This suggests that incremental increases in the stigmatized minority Catholic population are driving Belfast’s interethnic violence, and this is consistent with a theory of group threat via exposure. It is further corroborated by qualitative accounts tracing Belfast’s earliest reported ethnic tensions to the industrial migration of rural Catholics to working class neighborhoods that were almost completely Protestant before (O’Hearn, 1983).

In the appendix, we undertake several robustness checks. First, we test whether our results might be driven by deaths in the 1980s or 1990s, as might occur if ethnic population patterns in later decades diverged dramatically from those early in the conflict (and from the distribution of churches built prior to the conflict). Table A3 shows that the association between church diversity and violence is unchanged if deaths after 1980 or 1990 are excluded from the analysis. Church diversity’s association with violence remains very consistent, even if both the 1980s and 1990s are omitted. Deaths peak in the 0.55–0.60 range of diversity for the 1.5 km bandwidth, in the 0.63–0.67 range for the 2 km bandwidth, and in the 0.75–0.81 range for the 3 km bandwidth. Table A4 shows that the conflict-intensifying effect we observe using our church diversity data (Panel A) stands in contrast to a null effect obtained using the 1971 census data (Panel B), whether one looks at deaths in the 1970s or the 1980s–1990s. The usefulness of our estimation method is therefore evident whether looking at early-conflict or late-conflict deaths. Table A5 excludes victims whose precise locations of death could not be determined. Our CAIN data

Figure 5. Marginal effects of church diversity on violence using negative binomial specification with SES control. Histograms represent data density by level of church diversity variable.
Figure 6. Predicted violence, by 10 percent bins of Catholic or Protestant proportion of church diversity, by church density bandwidth.
contain 224 deaths geolocated to where the victims’ bodies were found, although there were no witnesses to their deaths and there is some chance that the bodies were moved prior to discovery. Our results are qualitatively unchanged when excluding these observations. Table A6 separately analyzes intracommunal killings—for instance, Catholic paramilitaries killing Catholic civilians suspected of being police informants. This type of violence likely follows a different logic from that of interethnic conflict, potentially confounding our estimate of diversity’s effect. Our main results are qualitatively unchanged by excluding intracommunal killings. Finally, Table A7 shows that our results are unchanged when using the Ethnolinguistic Fractionalization (ELF) index to construct our diversity variable.

4. Conclusion

To date, social scientists have had difficulty characterizing diversity’s effect on conflict. Results from rich literatures on conflict and on intergroup prejudice provide arguments for either a positive or a negative relationship. As a result, the verdict remains a largely empirical endeavor. Yet empirical analyses face their own set of challenges. In this paper, we identify a substantial impediment to ascertaining diversity’s true effect on conflict intensity: the fact that conflict generates migration and measurement error. Any research design that does not account for conflict-induced migration and measurement error leads to biased estimates, limiting our ability to draw valid inferences about the relationship between diversity and violence.

In the face of these challenges, we use the data-rich context of the Troubles in Northern Ireland to develop a new corrective approach, given the tendency of physical structures to reflect underlying demographic patterns there. We introduce a new geolocated church dataset and offer an analysis that is not mired by the measurement error and resulting bias that conflict and migration create. This analysis uncovers two key empirical results: the relationship between diversity and violence is curvilinear, with the strongest effects occurring at low-to-mid levels of diversity; and incremental increases in the Catholic proportion of church density are most associated with increased violence. Together, these results are consistent with a story of violence through group-threat exposure—a narrative further corroborated by qualitative accounts of the historical relationship between Catholics and Protestants in Northern Ireland. Indeed, as others have noted, Protestants were quite accommodating of Catholics when they represented only a small proportion of Belfast’s population; they grew increasingly suspicious and hostile as more Catholics moved into the city, threatening the Protestants’ dominant demographic and social position (Boal, 1982; O’Hearn, 1983).

We developed our solution in the context of Belfast, a single city, but our method has the broader potential to capture population demographics in other cases as long as two conditions hold. First, ethnic communities must construct or utilize different physical structures. Second, the location of the structures must correlate with settlement patterns. Under these conditions, the distribution of physical structures at conflict onset can serve as a useful proxy for the distribution of demographic groups, without the potential biases created by conflict migration and intimidation.

For instance, our method might shed light on settlement patterns and conflict dynamics in the Sri Lankan civil war, which began with deadly ethnic pogroms that displaced tens of thousands of minority Tamils (Mampilly, 2011; Weiss, 2012). The civil war rendered census enumeration impossible for decades, and prewar censuses were compromised by the undercounting of Tamils, who in many cases changed their names to pass as ethnic majority Sinhalese (Silva, 1999, 2007). However, religious structures might be used to uncover the population distributions of the Tamil (Hindu) and Sinhalese (Buddhist) ethnic groups prior to the outbreak of conflict. The temples corresponding to each group serve not only as sites of religious observance and ethnic identity formation, but also as markers of ownership, demonstrating the group’s historical presence and its claim to a particular parcel of land (Bastin, 2005; Schut et al., 2008; Harris,
Our method might also be used to reconstruct settlement patterns in countries where censuses habitually undercount particular groups (Kopstein and Wittenberg, 2011), or in the roughly two-thirds of countries (e.g., Turkey, France) where censuses record no ethnicity data at all (Aktürk, 2012).

Additionally, our empirical approach might be extended to qualitatively different contexts, such as ideological wars or ethnic civil wars with conventional front lines. In conventional civil wars, belligerents strategically eliminate the opposing side’s civilian supporters from territory they conquer (Balcells, 2011). Using physical structures to measure prewar settlement patterns would enable research on behind-the-lines civilian victimization in conventional ethnic civil wars. It could also enable research on irregular ideological civil wars if the opposing sides habitually used different physical infrastructure, such as party offices or meeting spaces, which might serve as physical proxies for the prewar ideological composition of a given area.

Substantively, our research aligns with recent work showing that changes in diversity, rather than absolute levels of diversity, are most threatening (Hopkins, 2010; Newman and Velez, 2014; Cikara et al., 2020). Catholic migration to Belfast did not become contentious until the Catholic population share grew beyond its negligible pre-industrial levels. But even as our research affirms previous qualitative findings, it raises questions for further exploration: under what conditions might the threat of exposure become an opportunity for prejudice-reducing intergroup contact?

Finally, how do our findings extend to contexts with more than two competing ethnic groups? Fortunately, our empirical strategy, using the locations of buildings instead of flawed population data, does not depend on the number of groups, but on the existence of pre-conflict structures belonging specifically to each ethnic community. As we note in the paper, our approach is compatible with various operationalizations of diversity, including the RQ measure and the ELF measure we use as a robustness check. Thus, our findings are not dependent on the specific operationalization of diversity in a two-group case. However, other formal-theoretic analysis has noted that the operationalization of diversity (polarization versus fractionalization, for instance) may matter in cases with more than three groups, given the differences in how these indices are computed (Esteban and Ray, 2008). Our research opens avenues for an empirical investigation of this issue, assessing potential differences in diversity’s effect, using alternative operationalizations in cases with more than two groups, absent the confounding flaws of population data from conflict zones.

Acknowledgements. We are indebted to Andrea de Barros and Máirelise Robinson for excellent research assistance.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psrm.2022.36. To obtain replication material for this article, please visit https://doi.org/10.7910/DVN/56PIVI

References
Boal FW (n.d.) Forthcoming.
CAIN (2014) *Conflict and politics in Northern Ireland*. Derry/Londonderry, Northern Ireland: University of Ulster.
Cikara M, Fouka V and Tabellini M (2020) Hate crime increases with minoritized group rank. osf.io/2x3kw.
Fjelde H and Hultman L (1971) *Police will not be told*. Irish Independent, p. 15.
Lowe M (forthcoming) Types of contact: a field experiment on collaborative and adversarial caste integration.


Thangavelu V (2021) Gotabaya Rajapaksa government had opened another battle front against Hindu Tamils. *Colombo Telegraph*.


Weidmann NB (2011) Violence ‘from above’ or ‘from below’? The role of ethnicity in Bosnia’s Civil War. *Journal of Politics* 73, 1178–1190.


---