



THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■

Economic History Working Papers

No: 300

The Last Yugoslavs: Ethnic Diversity, National Identity, and Civil War

Leonard Kukic
Universidad Carlos III de Madrid

October 2019

The Last Yugoslavs: Ethnic Diversity, National Identity, and Civil War

Leonard Kukić*

Universidad Carlos III de Madrid, Clio III - Juan March Institute and Department of Social Sciences
e-mail: lkukic@clio.uc3m.es

August 18, 2019

Abstract

Nation-building is often proposed as a device for integration and conflict reduction in ethnically divided societies. This paper analyses the role of interethnic contact in the process of nation formation within the unique historical setting of the multi-ethnic Yugoslavia. Using historical border changes as a proxy for exogenous population movements that influenced ethnic diversity, I find that interethnic contact stimulated the formation of the Yugoslav national identity. In addition, aligned with the notion that nation formation can reduce the incidence of ethnic conflict, I find that areas with more self-declared Yugoslavs experienced a lower intensity of conflict during the Bosnian War of 1992-1995.

Keywords: National Identity; Ethnic Diversity; Civil War; Yugoslavia.

JEL Classification Numbers: N4; Z13; O10

*I thank my PhD supervisors at the LSE, Max-Stephan Schulze and Tamás Vonyó, for invaluable feedback and support. I also thank Sascha Becker, Christoph Koenig, James Fenske, Andy Ferrara and Nikolaus Wolf for their comments and suggestions. I benefited from presenting at the EUI, Split University, Groningen University, Swansea University, Cologne University, MIT, Warwick University, Princeton University, LSE, University of Oxford, Prague University of Economics, Humboldt University, and the Institute of Economics in Zagreb. All errors are mine.

1 Introduction

Ethnically divided societies tend to be characterised by a higher incidence of conflict, weaker institutions, and lower economic growth.¹ In this context, nation-building is often proposed as a device for integration in ethnically divided societies. Nevertheless, the determinants and implications of national sentiment remain imperfectly understood.² This paper analyses the role of interethnic contact in the process of nation formation.

The relationship between ethnic diversity and national sentiment is ambiguous.³ On the one hand, interethnic contact might breed hatred and conflict. The long history of anti-Semitism in Europe is an evident example (Voigtländer and Voth, 2012; Becker and Pascali, 2019; Koyama et al., 2019). On the other hand, psychologists have long recognised that contact with members of another ethnic group can lead to a reduction of prejudice and negative sentiment (Allport, 1954). The impact of interethnic contact might also depend on the relative size of different groups (Esteban and Ray, 2011; Esteban et al., 2012), or on the physical proximity between them (Montalvo and Reynal-Querol, 2017). Whether intergroup contact is conducive to integration or to conflict is therefore an empirical issue.

Yugoslavia provides a unique historical setting to study the relationship between ethnic diversity and national identity. Although the formation of the Yugoslav identity failed because the country dissolved, there were people who felt Yugoslav. Some people chose to identify with the wider Yugoslav community, rather than with some of the narrower ethnic communities.⁴ This is empirically useful, because it allows me to exploit the variation in national self-identification across the country.

Moreover, Yugoslavia was extremely diverse. There was a great deal of variation in economic, cultural and historical conditions. As such, Yugoslavs were fond of describing their country as one with two alphabets, three religions, four languages and five nations. This high diversity is empirically useful as well, because it allows me to exploit the variation in socio-economic conditions.

I analyse the relationship between ethnic diversity and Yugoslav identity at municipal-level. I focus the analysis on the last population census of 1991, conducted just before the disintegration of the country and the outbreak of violence. I analyse a sample of 433 Yugoslav municipalities, and measure the strength of Yugoslav identification by the fraction of people who answered “Yugoslav” to the open-ended nationality question of the census. I measure ethnic diversity by the ethnic fractionalisation index, under which many small groups will result in a high value.

¹See Alesina and La Ferrara (2005) for a literature survey, and Esteban and Ray (2011); Esteban et al. (2012) for the theory and the empirics underlying the various measures of ethnic divisions.

²Seminal qualitative works that study the topic include Weber (1976); Anderson (1983); Gellner (1983); Hobsbawm (1991).

³I use the terms “interethnic contact” and “ethnic diversity” synonymously. I also use the terms “national identity”, “national sentiment” and “national feeling” synonymously.

⁴Yugoslav identity was different to an ethnic-centred identity. Scholars studying the topic argue that Yugoslav identity provided a “broader” national identity category for those who rejected the “narrower” ethnic identity category (Ramet, 1992; Wachtel, 1998; Sekulić, 2004). Yugoslav identity was multicultural and pan-national. It was based on the understanding of a nation as a cluster of “close” ethnic and cultural groups with equal rights, contrasting the mutually antagonistic identities of the ethnicities that formed the country (Sekulić, 2004). To illustrate this further; the distinction between Yugoslav and Slovene identity, for example, is comparatively similar to the distinction between British and English identity. The interested reader can see Colley (1992) for a historical analysis of English identity.

Maps in figure 1 suggest that ethnically diverse municipalities were associated with a greater share of people that identified themselves as Yugoslavs. Basic correlations confirm this. One way to interpret this is that interethnic contact stimulated Yugoslav sentiment. An alternative explanation is that Yugoslavs conglomerated to ethnically diverse areas that tend to be more tolerant. Alternatively, there may be other important factors, such as urbanisation and schooling, that are correlated with both ethnic diversity and Yugoslav sentiment. Moreover, my key variables might be mis-measured, biasing the estimates. I pursue several strategies to determine whether the uncovered correlations are, in fact, causal.

My first strategy is to control for various socio-economic factors that may be relevant, like economic development, but also for many other cultural and historical factors. I find that controlling for various observable characteristics has little effect on the estimated positive effect of ethnic diversity on national feeling.

My second strategy is to use the framework of [Altonji et al. \(2005\)](#) and [Bellows and Miguel \(2009\)](#). I calculate how much greater the influence of unobservable factors would need to be, relative to observable factors, to render my results insignificant. I find that the influence of unobservable factors would have to be about six times greater than observable factors. It is thus unlikely that unobserved heterogeneity can explain away my results.

My third strategy is to conduct the analysis at a hyper-local, settlement, level. Settlements presumably better correspond to the daily environment of an individual, and thus to the interethnic contact that individual actually experiences. Indeed, this is exactly what I find: the positive effect of ethnic diversity is much larger at a lower level of aggregation. Furthermore, once I control for ethnic polarisation, under which a few large groups will result in a high value, the effect of ethnic fractionalisation is much larger as well. Both factors suggest that my baseline results provide lower bound estimates of the positive relationship between ethnic diversity and national identification.

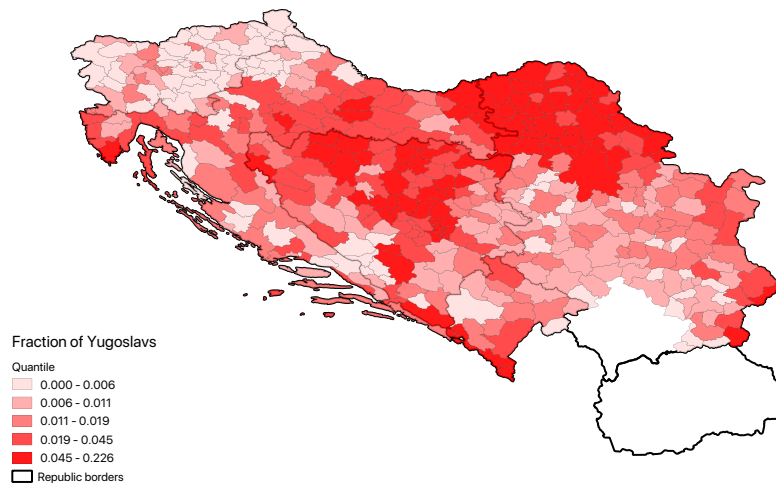
My final strategy to establish causality is to use a plausible source of exogenous variation in ethnic diversity. I would ideally want to find random shocks to population movements that influenced ethnic diversity, with these population movements having no direct impact on national sentiment. I can approach such a scenario by using border changes between the various empires that historically ruled and divided the lands of former Yugoslavia (e.g. Habsburgs, Ottomans and Venetians).

The plausibility of the instrument is based on three premises. First, there is a great deal of variation in historical border changes in Yugoslavia. Second, historians argue that border changes caused shocks to the ethnic composition of the areas that later formed Yugoslavia. Third, the impact of past shocks to ethnic composition persisted to the period of socialist Yugoslavia. I extensively document the historical plausibility of the instrument in section 4.1. To improve its plausibility, I measure border changes during the premodern and early-modern era (15th of 19th centuries). It is unlikely that border changes were endogenous to ethnic diversity during the premodern era, given that national identities are modern phenomena that emerged in the 19th century ([Weber, 1976](#); [Anderson, 1983](#); [Gellner, 1983](#); [Hobsbawm, 1991](#)).

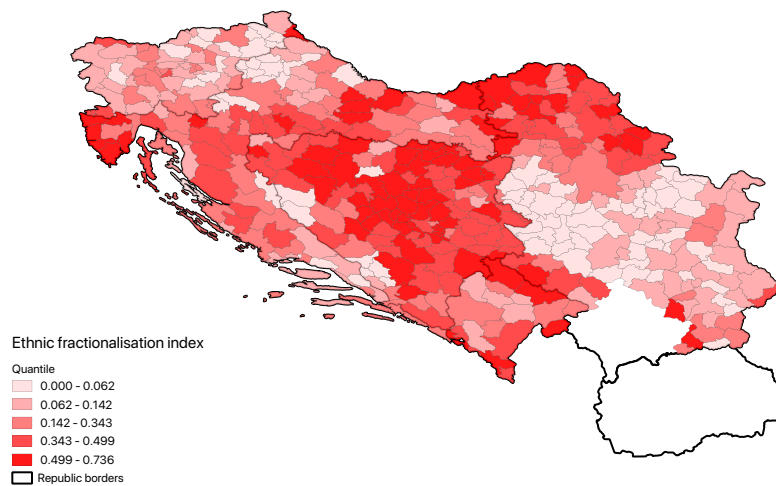
The instrumental variable (IV) estimates imply that moving from perfect ethnic homogeneity (zero diversity) to the sample average of 0.28 increases the fraction of Yugoslavs by 3.6 percentage

Figure 1: Population share of self-declared Yugoslavs and ethnic fractionalisation in Yugoslav municipalities, 1991

(a) Population share of self-declared Yugoslavs



(b) Ethnic fractionalisation index



Notes: The correlation coefficient between the share of Yugoslavs and ethnic fractionalisation is 0.56. I analyse a sample of 433 Yugoslav municipalities, which form 87 per cent of the total amount of municipalities. I exclude Macedonia and Kosovo from the analysis due to data issues. Kosovar Albanians boycotted the 1991 population census due to Serbia's dissolution of Kosovo's autonomy, while the 1991 Macedonian census was uncompleted due to ethnic tensions between ethnic Macedonians and Albanians. The data and the data sources are explained later in the text.

points, relative to the average of 0.027. To address the concern of whether the exclusion restriction is satisfied, I perform several robustness checks.

I begin by conducting a placebo test. I examine the reduced-form relationship between border changes and Yugoslav sentiment before and after the dissolution of Yugoslavia. Border changes after the dissolution of Yugoslavia could not have impacted ethnic diversity before the dissolution of the country, and this is exactly what I find.

I also relax the exclusion restriction following the framework of [Conley et al. \(2012\)](#). I find that the instrument would have to deviate very far from the exclusion restriction to render my estimates insignificant. Moreover, the IV results are robust to the inclusion of additional controls, an alternative measure of national sentiment, as well as to the different levels of data aggregation. While each of the empirical settings I use can be criticised on its own, the totality of evidence suggests that the IV is plausibly exogenous.

After establishing that ethnic diversity positively affected Yugoslav sentiment, I turn to the task of establishing a channel of causality. The analysis reveals that the key channel linking ethnic diversity and Yugoslav identification is ethnic intermarriage. Ethnic diversity influenced intermarriage, which stimulated Yugoslav sentiment. My interpretation is that the hybrid Yugoslav identity provided an alternative to forcing a single category to intermarried individuals and their children.

The second part of the paper documents some of the consequences of Yugoslav identity. I find that municipalities that contained a larger fraction of self-declared Yugoslavs were negatively associated with the intensity of casualties during the Bosnian War of 1992-1995. While it is difficult to establish causality, it seems plausible that individuals with a multicultural and pan-national sentiment were less likely to engage in an ethnic-based conflict, and die. Although Yugoslavs died less, they lost their country.

This paper contributes to a nascent literature within economics that analyses the causes and consequences of identity. In an influential set of articles, [Akerlof and Kranton \(2000, 2002, 2005\)](#) incorporate identity, a person's sense of self, into economic analysis. The empirical literature that studies the determinants of national identity follows their lead. [Blouin and Mukand \(2019\)](#) argue that media policy of the government influences national sentiment, while [Clots-Figueras and Masella \(2013\)](#); [Cinnirella and Schueler \(2018\)](#); [Bandiera et al. \(2019\)](#) argue that educational policy matters as well. Moreover, history can cast a long shadow, which is evident in the influence of colonial legacy on national identities in Africa ([Ali et al., 2018](#)). War might also matter - either through incentivising the state to provide public goods ([Alesina et al., 2017](#)), or through mobilizing the people against an external enemy ([Dell and Querubin, 2018](#)). This paper makes a contribution to the literature by analysing the role of interethnic contact in the process of nation formation.

This paper is furthermore related to a large literature that studies the relationship between ethnic diversity and development (see [Alesina and La Ferrara \(2005\)](#) for a survey). The contribution of the paper to this literature is twofold. First, it reveals one of the origins of ethnic diversity, similar to [Ahlerup and Olsson \(2012\)](#); [Michalopoulos \(2012\)](#), and [Ashraf and Galor \(2013\)](#). Second, if identity is part of the utility function as [Akerlof and Kranton \(2000, 2002, 2005\)](#) argue, this paper reveals additional economic importance of ethnic diversity. The results

imply that ethnic diversity might have an indirect impact on utility through influencing identity.

Finally, this paper is also related to panoply of research about the dissolution of Yugoslavia.⁵ A large body of literature stresses the role of rising ethnic nationalism in the disintegration of the country (see [Jurajda and Kovač \(2016\)](#) for a novel treatment). However, the determinants of the flip side of ethnic nationalism - Yugoslav identity - are largely unknown.

This remainder of the paper is organised as follows. Section 2 documents the historical background underlying the idea of a Yugoslav nation. It also provides a conceptual framework about the relationship between ethnic diversity and national identity. Section 3 provides a set of OLS estimates, alongside a set of robustness checks. Section 4 documents the IV estimates, where the plausibility and robustness of historical border changes as an instrument for ethnic diversity is extensively discussed. Section 5 presents evidence that intermarriage is the key channel connecting ethnic diversity and Yugoslav identification, while section 6 documents the association between national identification and ethnic conflict. The final section 7 concludes the paper.

2 Historical background and conceptual framework

I begin this section by documenting the historical background that underlined Yugoslav identity. I proceed by outlining a conceptual framework about the relationship between ethnic diversity and national integration. In appendix A.2, I provide a formal model to describe which individuals might adopt a Yugoslav identity. The literature that I report in this section largely inspires that model.

2.1 Historical background: Yugoslav identity

The idea of a Yugoslav nation emerged in the South Slav lands under the Habsburg rule during the mid-19th century. It was a branch of Pan-Slavism inspired by the ideals of national awakening of Romanticism and the French Revolution. “Yugo-Slavism” was based on the common cultural and linguistic characteristics of, primarily, the Croats, Serbs and the Slovenes. Due to cultural similarities, proponents of the Yugoslav idea believed that South Slavs belonged to a single Yugoslav nation. They believed that Croats, Serbs and the Slovenes were three tribes of one nation.

[Wachtel \(1998\)](#) argues that, in addition to cultural factors, political factors contributed to the attraction of a national synthesis. It seemed unlikely that any of these ethnicities could achieve autonomy within the Habsburg Empire on their own. Even if they managed to achieve political independence, it seemed very unlikely they could maintain it in the presence of neighbouring empires. For these reasons, cultural unity was the main aim of the movement, while political unification was left to an idealised distant future.

The distant future, however, arrived sooner than anticipated. The collapse of the Austro-Hungarian Empire in the wake of WWI delivered it, allowing the creation of a South Slavic state. The newly merged Kingdoms of Serbia and Montenegro joined together with most of the South Slavs under the Habsburg rule to form a Kingdom of Serbs, Croats and the Slovenes (later

⁵See [Jović \(2009\)](#) for a literature survey.

renamed to Yugoslavia). The South Slavs initially perceived the common state as an instrument of gradual amalgamation of the various peoples, without the domination of any ethnicity. Reality soon confronted these utopian pre-war ideals of national unity. There were significant cultural, social and economic differences among the ethnicities that formed Yugoslavia (Lampe, 2000).

Political fault lines were characterised by a struggle between the supporters of a unitary state and the supporters of a federation. Political tensions soon transformed into ethnic tensions, and the two became synonymous, as the supporters of the unitary state were mostly Serbs, while the supporters of a federation were mostly the other ethnicities. As the supporters of the unitary state soon prevailed, the non-Serbs perceived the Kingdom of Yugoslavia as an extension of the pre-WWI Serbian Kingdom (Banac, 1984; Jović, 2009). This had strong implications concerning the process of national building. Serbian domination contributed to the consolidation of identities of the other ethnicities, hampering Yugoslav national unity (Sekulić, 2004).

A weak sense of common identity contributed to the fast collapse of the country under the German invasion of 1941 (Wachtel, 1998). After leading the powerful antifascist struggle, the communists, and their leader Tito, seized power in 1945. The communists were firmly committed to the reconstruction of Yugoslavia, reorganising the country into a federation. Each federal republic approximated the areas inhabited by the major Slavic ethnic groups to placate the ethnicities that formed the country.⁶ By this, the communists attempted to reduce political and ethnic divisions, and distance themselves from the unitary, interwar, Yugoslavia.

The communists did not pursue nation-building policies. To Edvard Kardelj, the main ideologist of Yugoslavia's Communist Party, the idea of a Yugoslav nation was a product of capitalism and nationalism (Kardelj, 1979). The socialist state should not make nor deny nations, like "bourgeoisie" states do. Instead, Tito and the federal leadership hoped that industrialisation and modernisation would erode ethnic divisions, creating support for the common state (Sekulić, 2004). By this ambivalent attitude towards Yugoslav feeling, the communists sought to distance themselves from the interwar Yugoslavia, where central authorities in Belgrade suppressed ethnic boundaries (Banac, 1984).

There is another reason for the absence of nation-building policies. The ideological consensus among the communist elite was that the state should be gradually weakened during the transition process to the communist utopia. It was based on the Marxist notion that the state should "wither away" (Jović, 2009). The "withering" of the Yugoslav state involved the establishment of labour-managed firms and the devolution of power to the federal units. Kardelj believed that these "socialist" features were the main unifying force of the country, rather than ethnic features (Jović, 2009).

"Yugoslav" was not an easy answer to give in response to the open-ended census question "What is your nationality?". The communist elites were disinclined to place Yugoslav identification on an equal footing with ethnic identification (Burg and Berbaum, 1989; Ramet, 1992; Hodson et al., 1994b; Wachtel, 1998).⁷ The scepticism towards Yugoslav identification was reflected

⁶Ordered by their size, these ethnic groups were Serbs, Croats, Slovenes, Bosniaks (referred to as Muslims during the existence of the country), Macedonians and Montenegrins. The two largest non-Slavic groups were Albanians, which were concentrated in the Serb province of Kosovo, and Hungarians.

⁷Political and politically influential positions were distributed by ethnic criteria, especially at the federal level (Banac, 1984). An individual lost access to these positions by publicly declaring herself as Yugoslav.

in population censuses. Those that declared themselves as Yugoslavs were placed under the category “Yugoslav - no national affiliation”. The implication was that Yugoslav identification was legitimate, although by this the citizen did not declare membership in an existing nation. The tendency to view Yugoslav identification as something lesser to ethnic identification stands in sharp contrast to the preceding Kingdom of Yugoslavia. During the interwar Yugoslavia, population censuses did not contain a nationality question. The central government treated everyone as a member of a single Yugoslav nation.

For these reasons, it was somewhat surprising that the share of population that declared themselves as Yugoslavs increased over time in population censuses. The first time the category Yugoslav appeared in population census was in 1961. Then, 1.7 per cent of the population declared themselves as Yugoslavs. By the 1981 census, the percentage of Yugoslavs increased to 5.4. Of all the Yugoslav republics, Croatia contained the largest number of Yugoslavs in its population, 8.2 per cent.

Republican leaderships were highly concerned about the increase of Yugoslav sentiment. They perceived it as the deliberate attempt by the federal authorities, or other republics, to shift the popular loyalties of their citizens, and erode their power base. Serbian nationalists often interpreted Yugoslav identification as the deliberate attempt by the Croatian leadership to decrease the Serb minority in their republic. Croatian nationalists made the same arguments on behalf of Croats in Serbia.

Sociological surveys indicate that Yugoslav identification increased further during the 1980s ([Hodson et al., 1994b,a](#)). With the rise of ethnic nationalism, the population share of Yugoslavs in the 1991 census sharply decreased to about 2 per cent. These were the last Yugoslavs that still believed in a common nation.

2.2 Conceptual framework

There are several reasons why ethnic diversity could have influenced Yugoslav sentiment. The first reason is that contact with members of another ethnic group can lead to a reduction of prejudice and negative sentiment. [Allport \(1954\)](#) united the early research of psychologists and social psychologists and provided the most influential exposition of the contact theory. He claimed that prejudice and hatred is a result of generalizations made about members of another group based on incomplete or mistaken information. It follows that prejudice can decline as one learns more about the others. This implies that interethnic contact might decrease ethnic antagonism and stimulate national integration.

There is some evidence in economics that validates the contact theory. [Boisjoly et al. \(2006\)](#) investigate the consequences of intergroup interactions in the setting of a university. They find that white students who were randomly assigned an African American roommate developed more positive attitudes towards the African Americans. [Ferrara \(2018\)](#) finds that workplace diversity in the 1960s, induced by the high causality rate among the American whites during WWII, stimulated racial integration in the U.S.. In another historical study, [Jha \(2013\)](#) finds that medieval trading ports in India, characterised by intense interethnic contact, had a higher degree of ethnic tolerance between the Muslims and Hindus.

The second reason why ethnic diversity can influence national sentiment is because it can

decrease the cost of identity switching. [Caselli and Coleman \(2013\)](#) and [Manning and Roy \(2010\)](#) argue that ethnic identities are endogenous and that they can be switched, as long they are not based on strong ethnic markers. Nevertheless, changing an identity entails a cost. Such costs could include monetary costs, like lost access to jobs allocated by ethnic criteria, or physic costs, like ostracism from an ethnic-based community. A higher ethnic diversity might decrease the number of individuals that could punish another individual for “betraying” her ethnic roots, decreasing the cost of identity switching, and stimulating national sentiment.

Finally, ethnic diversity can influence national integration through stimulating ethnic intermarriage. Sociologists have long recognised that intermarriage is of central importance in interethnic relations. Intermarriage weakens the delineation of ethnic boundaries, and thereby decreases the salience of ethnic identities ([Davis, 1991](#); [Qian and Lichter, 2007](#)). Moreover, intermarried individuals, and especially their children, are less likely to identify with a single ethnicity ([Waters, 1990](#); [Xie and Goyette, 1997](#)). Both reasons suggest that identification with a broader nationality might provide an alternative to forcing a single ethnic choice to intermarried individuals and their children.

3 OLS estimates

This section provides evidence on the relationship between ethnic diversity and national sentiment. I start by outlining the data, and providing a set of OLS regressions. I subsequently conduct an array of robustness checks that lend some credibility to the OLS correlations. The results are robust to a sub-sample analysis, different levels of aggregation (settlement and individual level data), ethnic polarisation, and selection on un-observables relative to the selection on observables.

3.1 Data

Due to the decentralised nature of the country, Yugoslavia offers a rich set of socio-economic indicators at a local level. Although I rely on a variety of data sources, the 1991 population census is the main data source that I use. The Yugoslav statistical office conducted the census, but the results were subsequently published by the statistical offices of the successor states of Yugoslavia. [Table 1](#) provides descriptive statistics of the main data I use in this paper. I describe the data in detail with their sources in [appendix A.1](#).

The main dependent variable is the fraction of population that answered “Yugoslav” to the open-ended nationality question of the census. The main independent variable of interest is ethnic diversity, measured by the ethnic fractionalisation index, e :

$$e = 1 - \sum_{i=1}^M s_i^2 \quad (1)$$

where s is the percentage share of an ethnicity in the population of municipality i . The index measures the probability that two randomly drawn individuals from a population belong to two different ethnic groups ([Alesina et al., 2003](#)). Its maximum is reached when each person in a given population belongs to different ethnic groups (value of one). I exclude people that

Table 1: Descriptive statistics in socialist Yugoslavia and beyond

	Observations (1)	Mean (2)	Standard deviation (3)	Minimum (4)	Maximum (5)
<i>Main outcome variables</i>					
Fraction of Yugoslavs	434	0.027	0.034	0.000	0.226
Population share of war casualties, 1992-1995	109	0.024	0.024	0.002	0.198
<i>Main explanatory variables</i>					
Ethnic fractionalisation (index)	434	0.276	0.210	0.004	0.736
Intermarriage (ratio)	109	0.071	0.054	0.000	0.276
Border changes, 1421-1816	434	1.592	0.920	0.000	5.000
<i>Control variables</i>					
Output p.c.	434	8.611	5.527	0.908	58.905
Public goods exp. p.c.	434	1.458	4.891	0.055	98.389
Population density	434	125.997	340.657	6.359	5270.200
Avg. years of schooling	434	7.811	1.174	5.237	11.643
Social sector labour (population fraction)	434	0.596	0.220	0.064	1.000
Youth labour actions (population fraction)	434	0.050	0.084	0.000	0.725
WWII partisan veterans (population fraction)	434	0.006	0.005	0.001	0.056
Fascist terror (population fraction)	434	0.031	0.040	0.000	0.450
1980s' generation (population fraction)	434	0.225	0.026	0.126	0.299
Working-age population (population fraction)	434	0.668	0.030	0.379	0.727
Religious fractionalisation (index)	434	0.259	0.212	0.003	0.718
Ethnic polarisation (index)	434	0.455	0.315	0.008	0.985
Minorities (population fraction)	434	0.295	0.300	0.005	1.000
Muslims (population fraction)	434	0.125	0.228	0.000	0.975
Duration of Habsburg rule (years)	434	160.546	173.357	0.000	497.000
Agricultural (wheat) suitability	434	2684.629	229.498	1792.667	3022.143
Terrain roughness (index)	434	20.632	12.223	1.706	88.696
Longitude	434	18.109	2.356	13.625	22.823
Latitude	434	44.597	1.106	42.000	46.716

Note: The monetary variables are expressed in Yugoslav Dinars (at current prices, in millions). The data and the data sources are described in detail in appendix A.1.

declared themselves as Yugoslavs from the ethnic fractionalisation index because the econometric exercises would otherwise suffer from reverse causality.

3.2 Baseline OLS estimates

I now use ordinary least-squares (OLS) regressions to examine the relationship between Yugoslav identification and ethnic diversity.⁸ I extensively discuss the rationale underlying the choice of each control variable that I use. The linear regressions are of the following form:

$$y_i = \alpha + \beta e_i + \gamma X_i + \epsilon_i \quad (2)$$

where y_i is the fraction of Yugoslavs in population of municipality i , e_i is ethnic fractionalisation, X_i is a set of controls, and ϵ_i is a random error term. The coefficient of interest is β , the effect of ethnic diversity on Yugoslav identity.

Table 2 reports the OLS regressions of the fraction of Yugoslavs against ethnic fractionalisation for a variety of specifications. Column 1 indicates that there is strong positive correlation between ethnic fractionalisation and Yugoslav sentiment. The coefficient is economically substantial. It implies that moving from zero ethnic diversity to the sample mean of 0.28 is associated with an increase in the fraction of Yugoslavs by 2.5 percentage points (0.09×0.28), relative to an overall

⁸Appendix A.3 experiments with different estimators. They have no meaningful bearing on the results.

sample mean of 0.027. The effect is statistically significant at the 1% level, whereby standard errors throughout the main text are robust and clustered at municipal level.⁹

Columns 2-14 add covariates that might have impacted Yugoslav identification. I add the covariates sequentially in order to assess the stability of the coefficient on ethnic diversity. If the coefficient on ethnic diversity is unstable under different specifications, this might indicate in an informal way the presence of multicollinearity and/or the omitted variable bias.

Social scientists argue that the formation of nation-states is historically associated with modernisation and economic development (Weber, 1976; Anderson, 1983; Gellner, 1983; Hobsbawm, 1991). For that matter, as stated in section 2, the communist elite of Yugoslavia hoped that economic development would diminish the salience of ethnic identities as a shared prosperous community would emerge. To control for economic development, column 2 adds output per capita, column 3 adds population density as a proxy for the urbanisation rate, and column 4 adds average years of schooling.

Symbolic representations of a new state, rituals of history, and images of shared destiny, are important elements of national identity (Weber, 1976; Anderson, 1983; Gellner, 1983; Hobsbawm, 1991). One of the symbolic representations of the socialist Yugoslav society was the labour-managed firm (Burg, 1983). Column 5 includes the fraction of labour that was employed in labour-managed firms, or in the so-called “social sector”.

Youth labour brigades were an important tool in the development of socialist identity. Two million Yugoslavs participated in these brigades between 1942 and 1990 (Popović, 2010). Youth labour actions involved mass-scale voluntary labour implementing various projects. The most famous one was the “Brotherhood and Unity” motorway that connected Zagreb and Belgrade. Given that youth labour brigades were multi-ethnic, the communist party also used them as a mechanism to create interethnic ties of mutual acceptance (Popović, 2010). Column 6 includes the fraction of population engaged in youth labour actions.

Images or memory of a shared past could have further inspired Yugoslav sentiment. Connor (1984) argues that the anti-fascist struggle during WWII, led by the communists, inspired a sense of belonging to Yugoslavia. The antifascist struggle, for example, took a central place in the study of history in primary and secondary schools. The communists deliberately emphasized the interethnic nature of that struggle to heal ethnic divisions and to create support for the common state (Koren, 2012). As such, column 7 includes the population share of WWII partisan veterans to control for anti-fascism and the cultural transmission of its memory in the style of Bisin and Verdier (2001). Additionally, column 8 controls for monuments to anti-fascism, which contains sites that were officially considered to be of high importance in defeating the fascists during WWII.¹⁰ Finally, column 9 includes the population share of individuals exposed to fascist terror during WWII (e.g. survivors of concentration camps, forced labour, jails and war captivity).

Provision of public goods is a major legitimating tool of states, stimulating national unity (Weber, 1976; Alesina et al., 2017). Column 10 adds expenditure on public goods per capita,

⁹Appendix A.11 clusters standard errors at regional-level to allow for the correlation between the error terms in observations within the regions. Clustering standard errors at a different level keeps the key results unchanged.

¹⁰These sites, for example, include the municipality of Drvar, which was the location of the unsuccessful Nazi airdrop designed to assassinate Tito in partisan headquarters in May of 1944.

Table 2: OLS estimates, dependent variable: fraction of Yugoslavs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Ethnic fractionalisation	0.0913*** (0.008)	0.0914*** (0.008)	0.0892*** (0.008)	0.0887*** (0.007)	0.0884*** (0.008)	0.0911*** (0.008)	0.0908*** (0.008)	0.0906*** (0.008)	0.0920*** (0.008)	0.0912*** (0.008)	0.0956*** (0.008)	0.0939*** (0.009)	0.0885*** (0.007)	0.0964*** (0.009)	0.0794*** (0.008)
Output p.c.		0.0002 (0.000)													-0.0001 (0.000)
Population density			0.0000 (0.000)												0.0000** (0.000)
Avg. years of schooling															0.0000 (0.000)
Social sector labour				0.0037*** (0.001)											0.0153*** (0.002)
Youth labour actions					0.0143** (0.006)										-0.0057 (0.010)
WWII partisan veterans						0.0097 (0.013)									-0.0195* (0.010)
Monuments to anti-fascism							0.5712* (0.332)								0.3418 (0.468)
Fascist terror								0.0059* (0.003)							-0.0033 (0.003)
Public goods exp. p.c.									0.0359 (0.025)						0.0518** (0.025)
Federal aid										0.0001 (0.000)					-0.0011*** (0.000)
1980's generation											-0.0114*** (0.003)				-0.0043 (0.003)
Army presence												-0.0595 (0.071)			-0.1406 (0.091)
Region effects	No	No	No	No	No	No	No	No	No	No	No	No	0.0144*** (0.003)	Yes	0.0069** (0.003)
Observations	434	434	434	434	434	434	434	434	434	434	434	434	434	434	434
R-squared	0.319	0.319	0.343	0.335	0.327	0.319	0.325	0.325	0.320	0.319	0.346	0.320	0.355	0.381	0.555

Notes: The data and the data sources are described in detail in appendix A.1. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

which was locally funded and provided. Column 11 adds a dummy variable whether a municipality was recipient of federal aid.

Ramet (1992) argues that Yugoslav sentiment during the late 1970s and the 1980s was promoted by the emergence of a strong pan-Yugoslav rock scene oriented towards the youth. Column 12 includes the fraction of population that was young during the early 1980s, when the rock scene was particularly vibrant (Ramet, 1992).

Burg and Berbaum (1989) argue that Yugoslav identification signified political integration and regime support. The Yugoslav army, as the largest beneficiary of federal expenditure, had a vested interest in the survival of Yugoslavia. Dimitrijević (2001) argues that the army strongly promoted loyalty to Yugoslavia during the obligatory service. Column 13 controls for army presence - sites of battalions and divisions, army bases, or facilities.

To test further the interpretation of Yugoslav identification as signifying regime support, in appendix A.4, I control for the population fraction of the Communist Party members. I restrict the analysis to Croatia due to data availability issues.¹¹ The coefficient on ethnic diversity remains significant, and is similar to those documented in table 2, which includes a larger Yugoslav sample. Conditional on ethnic diversity and the baseline controls, there is no evidence that the communists were more likely to feel Yugoslav.¹²

Finally, column 14 includes region effects to control for region-specific un-observables (Bosnia, Croatia, Montenegro, Slovenia and Serbia). Such un-observables could include educational policies, which were region-specific (Sekulić, 2004).

Throughout the specifications, the coefficient on ethnic fractionalisation remains highly stable. Column 15 includes jointly all the considered controls. Given the inclusion of region effects, column 15 specification exploits the within-regional variation in Yugoslav identification, holding constant regional institutions and policies. The analysis is thus not hampered by institutional heterogeneity, which could otherwise make it difficult to disentangle the effect of ethnic diversity from institutions.

The coefficient on ethnic fractionalisation in column 15, which includes all the controls, is very similar to that in column 1.¹³ This is reassuring, as it indicates in an informal way that my baseline OLS specification is unlikely to suffer from a serious form of an omitted variable bias. Nevertheless, I address further the issue of omitted variable bias later in the text.

3.3 Ethnic polarisation

In this paper, I measure ethnic diversity by the ethnic fractionalisation index. I can extract corroborative evidence by analysing whether the effect of interethnic contact persists when using a different measure of it. The alternative to the ethnic fractionalisation index that the literature

¹¹I have not been able to find Communist Party membership records for the other regions in the archives.

¹²Appendix A.8 also includes Communist Party membership using individual-level survey data.

¹³The coefficients on population density, educational attainment, fascist terror, and army presence, are positive and statistically significant. This suggests that economic development, as well as the exposure to WWII fascist terror and the Yugoslav army, was positively associated with Yugoslav sentiment. The coefficient on youth labour actions and public goods expenditure p.c. is negative and statistically significant. The negative coefficient on youth labour actions is unexpected. This might be a statistical artefact, as the effect is statistically significant only at the 10 per cent level. The negative sing on public goods expenditure p.c. is expected. Municipalities and the regional units provided public goods. The consumers of public goods presumably associated the provision of public goods with regional units, rather than with Yugoslavia *per se*.

frequently uses is the ethnic polarisation index (Alesina et al., 2003).¹⁴ This index measures how far a population is from a bimodal distribution.¹⁵ Montalvo and Reynal-Querol (2005b,a); Esteban and Ray (2011); Esteban et al. (2012) argue that ethnic fractionalisation and ethnic polarisation indexes are closely related, but not perfectly. Many small groups of similar size will result in a high value of the ethnic fractionalisation index, while a few large groups will result in a high value of the ethnic polarisation index.

These differences are subtle, but important. Montalvo and Reynal-Querol (2005b,a) argue that divisions along a few large groups (ethnic polarisation) affect conflict, rather than divisions along many small groups (ethnic fractionalisation). The corollary of this finding is that ethnically polarised communities, characterised by a few large groups, should have a weaker belief in a shared multi-ethnic community than ethnically fractionalised communities, characterised by many small groups.

Table 3 demonstrates that this is indeed the case. In column 1, I first reprint the effect of ethnic diversity from table 2 for ease of comparison. In column 2, I replace ethnic fractionalisation by ethnic polarisation. The coefficient on ethnic polarisation is much weaker than the coefficient on ethnic fractionalisation. This implies that the effect of interethnic contact on national unity is weaker in communities characterised by a few large groups, relative to communities characterised by many small groups. In column 3, I include jointly the ethnic fractionalisation and polarisation indexes. Now, the coefficient on ethnic polarisation turns negative, while the coefficient on ethnic fractionalisation remains positive, but its size increases by about two times. Conditional on ethnic polarisation, moving from zero ethnic fractionalisation to the sample mean of 0.28 is associated with an increase in the fraction of Yugoslavs by 5.5 percentage points, relative to an overall sample mean of 0.027.

These findings imply two things. First, conditional on ethnic fractionalisation, ethnically polarised communities seem to reinforce the salience of ethnic identities, decreasing national sentiment. Second, if ethnic polarisation is omitted, the size of the coefficient on ethnic fractionalisation provides a lower bound estimate.

3.4 Sub-samples and survey data

As further corroborative evidence, I now analyse whether the relationship between ethnic diversity and national identity holds within several sub-samples. Given the inclusion of region effects, my baseline specification exploits the variation of variables within the regions. The obvious extension is to examine separately the data for each region that constituted Yugoslavia. In appendix A.6, I present a table showing the effect of ethnic diversity for the sub-samples. The coefficient on ethnic diversity in each region is broadly similar to the coefficient on ethnic diversity in table 2, which includes the full sample. Slovenia, however, is characterised by a significantly

¹⁴Another proxy for ethnic diversity that the literature sometimes uses is religious diversity (Alesina et al., 2003). Of course, religious diversity might matter independent of ethnic diversity. Appendix A.5 documents the relationship between religious diversity and national sentiment. Conditional on ethnic diversity, there is no evidence that religious diversity influenced Yugoslav self-identification.

¹⁵Following Montalvo and Reynal-Querol (2005a), the ethnic polarisation index, p , is defined as: $p = 4 \sum_{i=1}^N \sum_{j \neq i} s_i^2 s_j^2$, where s is the share of an ethnicity in a population.

Table 3: OLS estimates with ethnic polarisation, dependent variable: fraction of Yugoslavs

	(1)	(2)	(3)
Ethnic fractionalisation	0.0794*** (0.008)		0.1577*** (0.031)
Ethnic polarisation		0.0461*** (0.005)	-0.0563*** (0.021)
Baseline controls	Yes	Yes	Yes
Observations	434	434	434
R-squared	0.555	0.511	0.572

Notes: Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. The data and the data sources are described in detail in appendix A.1. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

lower coefficient on ethnic diversity. Apart for Montenegro, the coefficient on ethnic diversity is significant in all the regions.

In an ideal world, I would want to estimate the effect of ethnic diversity at individual-level. I can do this only for Croatia and Serbia. I derive individual-level data for each of these two regions from the 1990 pre-election surveys.¹⁶ These surveys include questions regarding the demographic and economic conditions of the respondents, including their nationality. I match this individual-level survey data with the ethnic diversity measure from the population census for the municipalities in which the respondents lived. This allows me to test the relationship between ethnic diversity and Yugoslav sentiment at individual-level. In appendix A.8, I show that individuals that lived in ethnically diverse municipalities were more likely to self-identify as Yugoslavs. Ethnic diversity and national sentiment are thus positively correlated even when combining different datasets.

3.5 Omitted variable bias and measurement issue

The results of the paper above demonstrate a strong positive correlation between ethnic diversity and Yugoslav sentiment. There are three reasons why this relationship should not be interpreted as casual. First, there is a problem of reverse causality, which should positively bias the OLS estimates. Yugoslavs might have conglomerated to ethnically diverse areas, or feel more comfortable in expressing themselves as Yugoslavs there, because such areas tend to be multicultural

¹⁶I was not able to find similar data for any other region.

and tolerant.

Second, even though I control for a range of observable factors, there are omitted factors that are correlated with both ethnic fractionalisation and Yugoslav identification. Such factors could include exposure to the idea of Yugoslav identity. As I cannot control for such factors, omitting them should positively bias the OLS estimates. I use the approach of [Altonji et al. \(2005\)](#) and [Bellows and Miguel \(2009\)](#) to assess the extent of the omitted variable bias. In appendix [A.9](#), following their approach, I find that selection on un-observables would have to be at least 6.7 times greater than selection on observables to overturn my results. This suggests that my baseline OLS specification is highly unlikely to suffer from a serious form of an omitted variable bias.

Third, if ethnic fractionalisation index is noisy, and does not correspond well to the intensity of interethnic contact, then the OLS estimates might be biased towards zero. Ethnic fractionalisation index will underestimate the impact of ethnic diversity if municipalities were ethnically segregated. For instance, within many municipalities, there was a village highly dominated by one ethnicity, next to a village highly dominated by another ethnicity. When I aggregate such villages into a municipality, this will show up as a high degree of ethnic diversity. This is potentially misleading, if the villagers from one settlement were not actually exposed to the villagers from the other settlement. Settlements (villages and town districts), rather than municipalities, presumably better correspond to the daily environment of an individual, and thus to the interethnic contact that individual experiences.

In appendix [A.7](#), I assess the extent of this measurement issue by conducting the analysis at a hyper-local, settlement, level. As expected, the coefficient on ethnic diversity is much higher at a lower level of aggregation. The estimates I present in the main text, which are based on municipalities, thus provide a lower bound effect of the interethnic contact.

Nevertheless, the issue of reverse causality remains. Furthermore, while the approach of [Altonji et al. \(2005\)](#) and [Bellows and Miguel \(2009\)](#) is useful in assessing the possible bound of the omitted variable bias, it cannot alleviate it. To address these problems, a plausible instrument for ethnic diversity is required. The instrument should be associated with the variation in ethnic diversity. It should have no direct impact on Yugoslav sentiment.

4 IV estimates

In this section, I use historical border changes as a plausible source of exogenous variation in ethnic diversity. The underlying premise of the instrument is that historical border changes are a reasonably good proxy for exogenous population movements that influenced ethnic diversity. I begin by outlining the historical plausibility of the instrument and conducting the baseline IV estimation. I proceed by using additional controls, conducting a placebo test, relaxing the exclusion restriction, and using an alternative dependent variable. I finish the section by assessing the remaining measurement issues.

4.1 Background: Border changes

There is a great deal of variation in Yugoslav lands concerning historical border changes. Areas of former Yugoslavia were previously divided by the Bosnian, Hungarian, Habsburg, Italian, Napoleonic, Ottoman, Serbian and Venetian rule. Scholars studying the region argue that border changes caused shocks to the ethnic composition of communities, primarily because they affected ethnic migration patterns (Valentić, 1990; Žerjavić, 1993; Jurković, 2005; Dugački, 2009).

Of course, until the late 19th century, ethnicities in the modern sense did not exist in the area of former Yugoslavia (Wachtel, 1998). However, border changes affected cultural diversity. Border changes implicitly affected ethnic diversity because ethnic identities in Yugoslavia emerged along religious and linguistic lines. Thus, for instance, the correlation coefficient between ethnic and religious fractionalisation index in Yugoslavia was 0.91.

Consider some historical examples as evidence. Because of the Ottoman invasion of modern-day Croatia during the 16th century, a significant fraction of the local population fled to modern-day Austria and Italy (Valentić, 1990). The emigrants were Slavic Catholics that today consider themselves Croatian. The descendants of these emigrants are known as *Burgenland* Croats in Austria and as *Molise* Croats in Italy.

Following Habsburg's reconquest of eastern Croatia from the Ottoman Empire in the 17th century, the Habsburgs established a military border as a bulwark against the Ottoman Empire. They colonised the area with Orthodox Christians (now Serbs) from the area of modern-day Bosnia. The Habsburgs offered some economic freedoms to the immigrants in exchange for service in the Imperial Army (Rothenberg, 1966). Furthermore, the landlords of some of the reconquered areas invited ethnicities from the other parts of the Habsburg Empire to reinvigorate the depopulated areas, like Czech and German speakers (Dugački, 2009).

To take another example, the recurrent wars between Venice and the Habsburg Empire led to frequent border changes in Istria and Dalmatia. Žerjavić (1993) argues that both the Venetians and the Habsburgs attempted to inhabit the depopulated settlements with immigrants. He reports that both rulers attempted to attract migrants by offering land, and fiscal and labour allowances. Naturally, under the Venetian rule, the migrants came from modern-day Italy. Under the Habsburg rule, the migrants came from the various parts of the multi-ethnic Habsburg Empire.

As a final example, consider Bosnia. Before the Ottoman invasions, Slavic Catholics (now Croats) and Slavic Orthodox Christians (now Serbs) inhabited the country. The long Ottoman rule of Bosnia (from mid-15th to late 19th century) stimulated a large fraction of the local Slavs to convert to Islam (Lampe, 2000). The converts eventually formed a distinct Bosniak identity, causing a change in the ethnic diversity of the region.

These historical episodes largely document examples where border changes caused positive shocks to the ethnic composition of communities. Nevertheless, I do not imply by the above selection of examples that this relationship should necessarily be positive across time and space. It could be negative in some sub-periods or regions.¹⁷ The rationale of the instrument rests on the observation that border changes caused shocks to ethnic diversity, whether positive or

¹⁷For instance, a sizeable Ottoman Muslim population fled Serbia as the country gained independence from the Ottoman Empire in the late 19th century (Pavlović, 2015).

negative.

I recognize that the underlying problem with the instrument is that it might be endogenous to ethnic diversity. For instance, the Adriatic city of Rijeka in modern-day Croatia became a Free State of Fiume in the aftermath of WWI. Due to its ethnic diversity, but also strategic location, various neighbouring powers laid ownership claims on Rijeka. Given the political standstill, the League of Nations elevated Rijeka to a status of an independent state as a sort of a compromise solution. The independent status of Rijeka was short-lived - Italy annexed it in 1924.

To improve the plausibility of the instrumental variable approach, I measure border changes during 1421-1816.¹⁸ It is unlikely that border changes were endogenous to ethnic diversity during the premodern and early modern era (1421-1816), given that national identities are modern phenomena that emerged in the 19th century (Anderson, 1983; Gellner, 1983; Hobsbawm, 1991). Thus, as Italy was unifying at the end of 19th century, the unification leader Massimo D'Azeglio famously stated: "We have made Italy, now we have to make Italians" (Stewart-Steinberg, 2007). Even in the case of France, with a long history of statehood, Weber (1976) argues that the French national identity became widely entrenched in the population only at the turn of the 20th century. In the case of Yugoslavia, which was a latecomer to the nationalist ideology, Sekulić (2004) argues that Croatian identity became widely entrenched only during the interwar period.

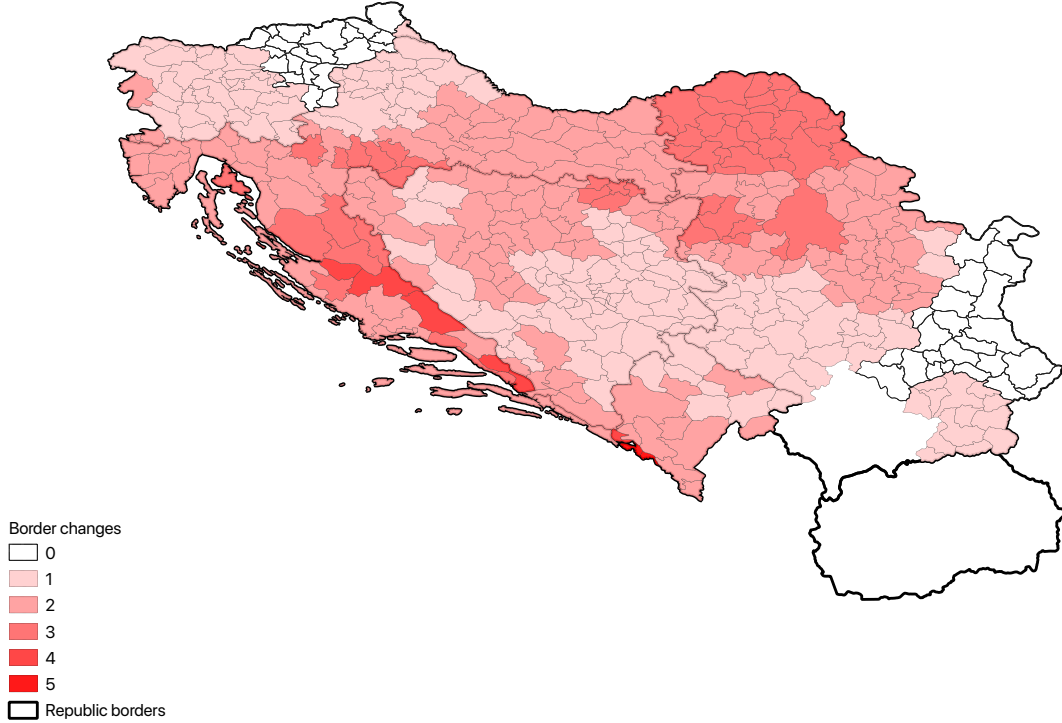
Nevertheless, although ethnic identities are modern phenomena, it might be that premodern and early modern border changes were endogenous to some religious or linguistic factors along which ethnic identities subsequently formed. History suggests that this is not the case. The three most important rulers driving the variation in border changes - the Habsburgs, Ottomans and the Venetians - presided over expansionist and imperialist empires. Historians argue that these empires were conquering areas according to strategic and economic criteria, rather than cultural criteria.

Murphey (1996) argues that material factors largely drove Ottoman conquests because soldiers and their commanders were entitled to booty. The material factor seems especially important in the Balkans. The largely autonomous *Bosnia Eyalet* led westward Ottoman conquests. It relied on irregular forces and mercenaries, which were particularly motivated by the prospect of booty. Economic considerations drove the Venetian military strategy as well. The Venetians targeted strategic coastal areas in order to protect their naval trade routes (Lane, 1973). The Habsburg military was driven primarily by economic criteria as well. They were attempting to secure their Austrian heartland against the Ottomans by expanding the frontier of its empire (Subtelny, 2007), with the aim of capturing the valuable coal and ore mines of Bosnia (Ruthner, 2019).

Although history suggests that this is not the case, it is, however, impossible to decisively reject the possibility that border changes do not satisfy the exclusion restriction. They might be correlated to the error term, or to some co-founders that influence national sentiment which I

¹⁸The initial year of the instrument, 1421, is taken to precede the Ottoman invasion of the West Balkans. The end year, 1816, includes the end of the Napoleonic presence in the region (Illyrian Provinces), while excluding border changes that might be related to the emergence of nationalism among the South Slavs at the end of the 19th century. For example, the 1878 Habsburg occupation of Bosnia and the 1878 establishment of the Kingdom of Serbia are likely both related to a rise of nationalism among the South Slavs. With the 1816 cut-off year, I exclude both of these 19th century border changes.

Figure 2: The amount of border changes in Yugoslav municipalities, 1421-1816



Notes: The instrument, border changes, is defined as the number of times a municipality was transferred between the various kingdoms, republics and empires that ruled and divided the lands of former Yugoslavia between 1421 and 1816. A border change equal to zero means that a municipality belong to single political entity during the sample-period. See appendix A.1 for details about the construction of the data.

do not control for. In section 4.3, I substantiate the presumption of the exclusion restriction by controlling for a range of additional factors that may be correlated with both border changes and Yugoslav sentiment. Furthermore, I conduct a placebo test in section 4.4, and in section 4.5, I relax the exclusion restriction and allow the instrument to have an impact on Yugoslav sentiment. All these tests suggest that the instrument is plausibly exogenous.

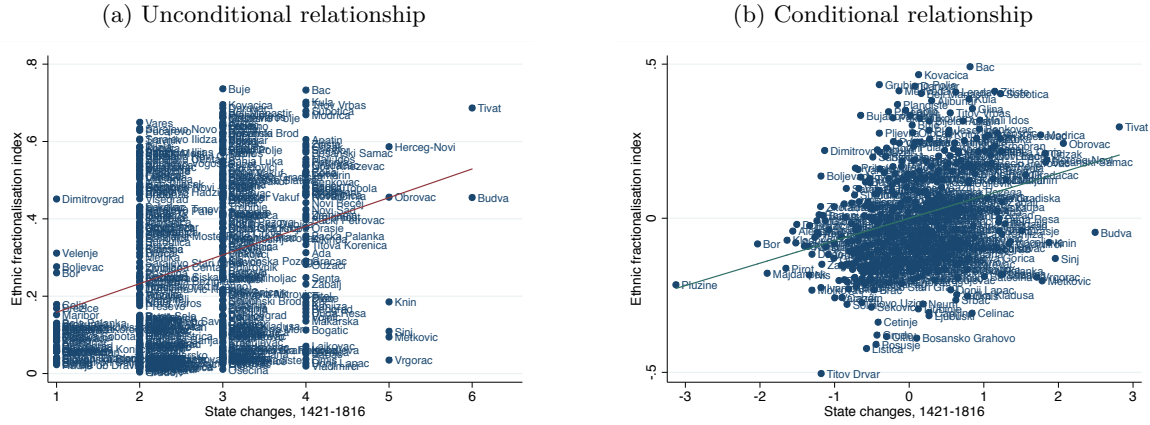
The instrument, border changes, is defined as the number of times a municipality was transferred between the various kingdoms, republics and empires that ruled and divided the lands of former Yugoslavia between 1421 and 1816. Figure 2 depicts border changes among Yugoslav municipalities. Figure 3.b plots the conditional relationship between ethnic fractionalisation and border changes, while figure 3.a plots the same relationship, but in a more transparent, unconditional, way. Both relationships are positive.

4.2 IV estimation results

Table 4 documents the two-stage least square (2SLS) estimates of equation 2. I treat ethnic fractionalisation, e_i , as endogenous, and estimate the following equation:

$$e_i = \lambda + \zeta s_i + \eta X_i + \nu_i \quad (3)$$

Figure 3: Relationship between ethnic diversity and border changes (1421-1816) in Yugoslavia



Note: The relationship between ethnic diversity and border changes in subfigure 3b is conditional on covariates included in column 15, table 2 specification.

where s_i is border changes during 1421-1816. The exclusion restriction is that s_i does not appear in equation 2. In the IV analysis presented in table 4, panel A documents the first stage estimate of ζ , the coefficient on border changes in equation 3. It implies that moving from zero border changes to its average of 1.6 increases ethnic diversity by 21 percentage points, relative to the sample mean of 0.28. The F-statistic, equal to 51.83, indicates a very strong first-stage relationship.

Panel B presents the second stage estimate, where the variation in ethnic fractionalisation caused by border changes predicts the fraction of Yugoslav. The coefficient is economically substantial. It implies that moving from perfect ethnic homogeneity (zero diversity) to the sample average increases the fraction of Yugoslavs by 3.6 percentage points, relative to the average of 0.027.¹⁹ The impact is significant at 1%.

In the appendix, to provide corroborative evidence, I experiment with alternative instruments (section A.10), and I apply the instrument to observations at settlement-level (section A.7) and at individual-level (section A.8). The results are reassuring - the effect of ethnic diversity persist even under alternative instruments and data.

4.3 Additional controls

The results presented in table 4 are contingent on the assumption that border changes had no direct impact on Yugoslav identification, conditional on the controls. This assumption is questionable if other factors correlated with border changes affected Yugoslav feeling. Despite already controlling for a rich set of observables, and despite using selection on observables to assess the bias from un-observables, in this section I include additional controls that may be correlated with both border changes and Yugoslav sentiment. I assess whether the inclusion of additional covariates affects the baseline results.

¹⁹2SLS coefficient on ethnic diversity is higher than the corresponding OLS estimate in column 15 of table 2. This suggests that the problem of measuring ethnic diversity dominates the issue of reverse causality and/or omitted factors. The OLS could also be a biased estimate of an average treatment effect that is different from the local average treatment effect estimated by the IV.

Table 4: 2SLS estimates

	(1)
Panel A: first stage outcome - ethnic fractionalisation	
Border changes, 1421-1816	0.0729*** (0.010)
Baseline controls	Yes
R-squared	0.403
Panel B: second stage outcome - fraction of Yugoslavs	
Ethnic fractionalisation	0.1353*** (0.027)
Baseline controls	Yes
Observations	434
Centered R-squared	0.476
Sanderson-Windmeijer first stage F-test	51.83***

Notes: Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. The data and the data sources are described in detail in appendix A.1. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Sekulić (2004) argues that minorities used Yugoslav identification to uplift them from their minority status. Column 1 of table 5 controls for the population share of minorities, which is related to the amount of border changes (section 4.1). The coefficient on minorities is insignificant and the coefficient on ethnic diversity remains robust. Conditional on ethnic diversity, there is no evidence that minorities were more inclined to feel Yugoslav.

The population share of Christians and Muslims might have influenced Yugoslav sentiment. The initial supporters of South Slav unity were exclusively Christians (Wachtel, 1998), which might had created path-dependency effects. Muslims, on the other, did not seem attracted to Yugoslavia before WWI (Bougarel, 2003). Moreover, the Habsburgs perceived themselves as the defenders of the Christian faith and civilisation against the Ottoman invasion (Subtelny, 2007). They might had thus conquered areas according to some religious factor, like population share of Christians or Muslims, although section 4.1 suggests this was not the case. Nevertheless, column 2 controls for religion (fraction of Muslims).²⁰ The coefficient on religious denomination

²⁰I do not control for Christians because a higher share of Muslims already implies a lower share of Christians.

Table 5: 2SLS estimates with additional controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: first stage outcome - ethnic fractionalisation									
Border changes, 1421-1816	0.0443*** (0.010)	0.0779*** (0.010)	0.0685*** (0.010)	0.0729*** (0.010)	0.0773*** (0.010)	0.0684*** (0.010)	0.0763*** (0.010)	0.0751*** (0.012)	0.0460*** (0.011)
Minorities	0.2957*** (0.047)								0.3026*** (0.054)
Muslims		0.2702*** (0.066)							0.3261*** (0.082)
Duration of Habsburg rule			0.0004*** (0.000)						0.0000 (0.000)
Agricultural (wheat) suitability				0.0001 (0.000)					0.0000 (0.000)
Terrain roughness					-0.0036*** (0.001)				-0.0003 (0.001)
Distance to coast						0.0007*** (0.000)			0.0011* (0.001)
Distance to river							-0.0007*** (0.000)		0.0003 (0.000)
Additional geographical controls	No	No	No	No	No	No	No	No	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.499	0.440	0.426	0.407	0.437	0.448	0.419	0.446	0.600
Panel B: second stage outcome - Fraction of Yugoslavs									
Ethnic fractionalisation	0.1587*** (0.051)	0.1266*** (0.024)	0.1296*** (0.028)	0.1353*** (0.027)	0.1279*** (0.023)	0.1335*** (0.029)	0.1383*** (0.026)	0.1555*** (0.033)	0.1368*** (0.045)
Minorities	-0.0177 (0.021)								-0.0116 (0.017)
Muslims		-0.0341*** (0.010)							-0.0374** (0.018)
Duration of Habsburg rule			0.0000** (0.000)						0.0001*** (0.000)
Agricultural (wheat) suitability				0.0000 (0.000)					0.0000 (0.000)
Terrain roughness					0.0004*** (0.000)				0.0007*** (0.000)
Distance to coast						0.0000 (0.000)			-0.0000 (0.000)
Distance to river							-0.0000 (0.000)		0.0000 (0.000)
Additional geographical controls	No	No	No	No	No	No	No	No	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	273	273	273	273	273	273	274	275	273
Centered R-squared	0.410	0.521	0.499	0.478	0.514	0.482	0.470	0.413	0.558
Sanderson-Windmeijer first stage F-test	18.77***	57.26***	46.03***	50.82***	59.65***	49.06***	56.45**	37.94***	16.03***

Notes: The variable Muslims is the ratio of Muslims to the population. Additional geographical controls are latitude, longitude, and an interaction between latitude and longitude. Baseline controls are output p.c., population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and regional effects. Baseline controls are identical to those in column 15 of table 2. The data and the data sources are described in detail in appendix A.1. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

is negative and significant, while that on ethnic diversity remains robust.

Historical affiliation to a political entity might have influenced Yugoslav sentiment in and of itself. [Becker et al. \(2016\)](#) find that historical Habsburg affiliation improves current trust in local public services. Moreover, [Grosfeld and Zhuravskaya \(2015\)](#) find that legacies of empires have an effect on political outcomes in Poland, while [Grosjean \(2011\)](#) finds that common empire affiliation improve cultural integration in eastern Europe. Historical memory of belonging to the multi-ethnic and multicultural Habsburg Empires might have thus influenced people to identify with the multi-ethnic and multicultural Yugoslav nation. Column 3 adds duration of Habsburg rule, which is correlated with the instrument, as longer Habsburg rule implies a lower amount of border change. The coefficient on Habsburg rule is significant and the coefficient on ethnic diversity remains robust.

Strategically and economically valuable areas - like areas with high agricultural suitability, that are close to coasts and rivers, and with a rough (or smooth) terrain - are attractive military targets ([Keegan, 1993](#)). Such areas might also have an advantage in trade and income generation, perhaps exposing them to a greater degree of interethnic contact, and affecting Yugoslav identification. Column 4 controls for agricultural (wheat) suitability, column 5 controls for terrain roughness, column 6 controls for distance to the coast, and column 7 controls for distance to the nearest river. Only the coefficient on terrain roughness is significant, while the coefficient on ethnic fractionalisation remains robust.

Given the spatial nature of the instrument, column 8 controls for latitude, longitude, and the interaction between latitude and longitude. The model thus controls for the location of municipalities. This is important, given that locations that are far from each other might be more different than locations that are close to each other, according to some factors I do not control for. Nevertheless, the coefficient on ethnic diversity remains robust with the inclusion of geographic coordinates.²¹

Column 9 includes jointly all the additional controls. Despite including variables that are presumably correlated with both border changes and Yugoslav sentiment, the results are highly reassuring. The coefficient on ethnic diversity (0.137) is nearly identical to the baseline coefficient in table 4 (0.135).

Including additional controls lends some credibility to the claim that the instrument is plausibility exogenous. Nevertheless, this exercise cannot decisively rule out endogeneity concerns, and further robustness checks seem warranted.

4.4 Placebo test

My IV strategy rests on the assumption that ethnic diversity is the only channel through which historical border changes influenced Yugoslav sentiment. If this assumption is correct, then a positive relationship between Yugoslav sentiment in 1991 and border changes that occurred later during the 1990s should not exist. To assess the validity of the IV estimates, I undertake this falsification test: I assess the relationship between border changes associated with the Yugoslav wars (1991-1995) and the strength of Yugoslav sentiment. As before, I measure Yugoslav

²¹ Appendix [A.11](#) experiments further with spatial models. My results are robust to the inclusion of a spatial lag of ethnic diversity.

sentiment by the nationality question of the 1991 population census, which took place before the outbreak of conflict and the dissolution of the country.

Table 6 reports the results. I begin by documenting in panel A the reduced-form relationship between border changes and Yugoslav sentiment. Column 1 depicts a strong positive reduced-form correlation between premodern border changes and Yugoslav sentiment. This correlation is consistent with the IV estimates (panels B and C) - historical border changes influenced ethnic diversity, and interethnic contact stimulated Yugoslav sentiment. In column 2, I conduct the falsification test. Border changes during the 1990s shows no systematic relationship to the Yugoslav sentiment in 1991. The reduced-form estimates are close to zero, and highly insignificant. This unsystematic relationship is consistent with the IV estimates, which are insignificant as well.

4.5 Relaxing the exclusion restriction

Although my falsification test provides evidence for the validity of the instrument, I recognize that the requirement of perfect exogeneity is a lingering problem. In order to gain a further sense of the robustness of the IV estimates, I relax the exclusion restriction that the instrument’s validity relies upon. Despite including a rich set of covariates, the validity of the instrument depends on perfect exogeneity which is, strictly speaking, unlikely to hold exactly. The instrument is rather plausibly exogenous.

I follow the approach of [Conley et al. \(2012\)](#) to relax the exclusion restriction. They propose a method which allows the instrument to have a direct impact on the outcome variable - in this case, independent of ethnic diversity effects. By varying the coefficient on the instrument, their method allows me to identify the threshold at which ethnic diversity becomes insignificant. In appendix A.13, I find that the impact of the instrument on Yugoslav sentiment would have to be equivalent to 87 per cent of the overall reduced-form effect to render the IV estimates insignificant. The analysis thus reveals that the instrument would have to deviate very far from the exclusion restriction to make my results insignificant.

4.6 Alternative dependent variable: Serbo-Croatian language

The totality of evidence so far suggests that the instrument is plausibly exogenous. In this section, I use an alternative dependent variable to corroborate further the existing evidence on the relationship between ethnic diversity and national identity. I use self-reported mother tongue as an alternative measure of national sentiment - social scientists have long perceived language as a critical component of national identity ([Weber, 1976](#); [Anderson, 1983](#); [Gellner, 1983](#); [Hobsbawm, 1991](#)).

The main official language of Yugoslavia was Serbo-Croatian. It is a pluricentric language with four, mutually intelligible, varieties - Bosnian, Croatian, Montenegrin and Serbian.²² The

²²Serbo-Croatian was chosen to be the standard language of the South Slavs in the 1850 “Vienna Literary Agreement” between the representatives of Croatia, Serbia and Slovenia ([Kordić, 2010](#)). The language was based on the *stokavian* dialect, which was the most widespread dialect spoken in the western Balkans. Before its choice as the standard language of the South Slavs, *stokavian* was spoken in parts of Croatia, Bosnia, Montenegro and Serbia. *Stokavian* thus seemed as a natural choice for a foundation on which a new standard language was to be built.

Table 6: Falsification tests, reduced-form relationships and IV estimates

	(1)	(2)
Panel A: OLS, reduced-form results		
Border changes, 1421-1816	0.0060*** (0.002)	
Border changes, 1991-1995		-0.0012 (0.004)
Baseline controls	Yes	Yes
R-squared	0.544	0.528
Panel B: 2SLS, first stage outcome - ethnic fractionalisation		
Border changes, 1421-1816	0.0640*** (0.020)	
Border changes, 1991-1995		0.0167 (0.033)
Baseline controls	Yes	Yes
R-squared	0.350	0.0632
Panel C: 2SLS, second stage outcome - fraction of Yugoslavs		
Ethnic fractionalisation	0.0938*** (0.025)	-0.0700 (0.319)
Baseline controls	Yes	Yes
Observations	211	211
Centered R-squared	0.6076	0.387
Sanderson-Windmeijer first stage F-test	10.10***	0.25

Notes: Panel A reports the reduced-form relationship between border changes and Yugoslav self-identification. I restrict the analysis to Croatia and Bosnia. They were the only two regions that experienced a variation in border changes within their territories during the 1990s. The falsification test is thus designed to go against my results, because including the excluded regions would further weaken the significance of the coefficients. Baseline controls are output p.c., population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. The data and the data sources are described in detail in appendix A.1. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

differences between these varieties of Serbo-Croatian are smaller than the differences in the varieties of German or English are (Kordić, 2010). All the speakers of Serbo-Croatian can understand each other perfectly well, irrespective of the exact variety of that language that they are using. Yet, the Yugoslav census allowed people to choose as their mother tongue the category “Serbo-Croatian” or, for example, the categories “Serbian” or “Croatian”.²³

The differences between these categories were subtle, yet important. Like elsewhere, language and nationalism were closely intertwined. For example, Croatian nationalists were often stressing the difference between the Croatian and the Serbian language in order to increase the salience of Croatian identity (Kordić, 2010). They perceived Serbo-Croatian as a cover to forcefully assimilate Croats into the wider Yugoslav community. The implication of choosing “Serbo-Croatian” as a mother tongue in the population census was therefore clear. It signified a sense of belonging to a wider Yugoslav community, while choosing “Croatian” as a mother tongue, signified a sense of belonging to a narrower Croatian community (Kordić, 2010).

In table 7, I use the fraction of population that reported Serbo-Croatian as their mother tongue as a proxy for the strength of Yugoslav feeling. The coefficient on ethnic diversity is positive and substantial - moving from zero ethnic diversity to the mean increases the fraction of Serbo-Croatian by 14.8 percentage points, relative to the average of 0.15 who reported Serbo-Croatian as their mother tongue. The effect is significant at 1%.

4.7 The meaning of Yugoslav identity

Although language provides additional evidence that ethnic diversity stimulated national feeling, there are three potential measurement issues that go against my results. The first issue is whether the Yugoslav category in the population census was a residual category. For example, it might be that stigmatised minorities, like Jews, choose Yugoslav identification to avoid stigmatisation. If so, Yugoslav identification did not reflect genuine Yugoslav feeling.

This seems very unlikely for two reasons. First, it is unclear why to Slovenian Jew for example, saying “Yugoslav” to the nationality question of the census taker would have been more convenient to avoid stigmatisation than saying “Slovenian”. Second, the Yugoslav population census was egalitarian and inclusive (Markowitz, 2007). For that matter, some acceptable answers were only remotely related to national identity. The census even reported regional identity as a legitimate form of a nationality, for example Dalmatian.

The second issue is whether the Yugoslav identification reflected ignorance and confusion. It might be that people of mixed ethnic background opted for Yugoslav identification to denote their mixed background, irrespective of any Yugoslav feeling. This is unlikely for two reasons. First, schooling is positively correlated with Yugoslav feeling (table 2). It is unlikely that people that identified themselves as Yugoslav were ignorant of the connotations of Yugoslav identity. Second, the population census allowed people to choose the “nationally undeclared” category. This category might have been a more obvious choice to people who wished to denote their mixed ethnic background, without choosing a specific identity category, alongside the associated connotations.

²³The census also allowed people to choose as their mother tongue Bosnian or Montenegrin, and many other varieties of Serbo-Croatian.

Table 7: 2SLS estimates with an alternative dependent variable - Serbo-Croatian language

(1)	
Panel A: first stage outcome - ethnic fractionalisation	
Border changes, 1421-1816	0.0729*** (0.010)
Baseline controls	Yes
R-squared	0.403
Panel B: second stage outcome - fraction of Serbo-Croatian speakers	
Ethnic fractionalisation	0.5368*** (0.132)
Baseline controls	Yes
Observations	253
Centered R-squared	0.531
Sanderson-Windmeijer first stage F-test	15.16***

Notes: The dependent variable is the fraction of population that reported Serbo-Croatian as their mother tongue in the 1991 population census. I exclude Serbia and Montenegro from the analysis. Their censuses did not report Serbo-Croatian as a separate language category, but grouped it instead in a single category with Serbian and Croatian. Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. The data and the data sources are described in detail in appendix A.1. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Finally, and perhaps most importantly, it is natural to wonder whether Yugoslav identity was based on substantive ideals and norms. If not, people might have identified themselves as Yugoslavs irrespective of their actual national feeling. In appendix A.15, I outline the meaning of Yugoslav identity, which was infused with rich content. I focus on the work and thoughts of the leading proponents of a supranational Yugoslav culture - the foundation of Yugoslav identity.

In case the reader finds the qualitative reasoning of this section unconvincing, in the remainder of space I provide quantitative evidence that Yugoslav identity signified national sentiment. I use individual-level data from a survey conducted in 1990 by a consortium of social science institutes across the country (Baćević, 1991). The aim of the study was to gauge the attitudes of Yugoslav citizens towards the socio-economic crisis that the country was experiencing at that time. In columns 1 and 2 of table 8, I include as dependent variables some of the answers to the survey question “What does the term “Yugo-Slavism” mean to you?”. The self-identified

Table 8: The meaning of Yugoslav identity

	(1)	(2)	(3)	(4)	(5)
	The meaning of Yugoslavism		Sense of belonging		
	Nation-formation	Supra-national identity	Locality	Region	Yugoslavia
Self-identified Yugoslavs	0.5604*** (0.211)	0.5559** (0.252)	-0.6785*** (0.111)	-0.8008*** (0.112)	0.3450*** (0.123)
Gender	Yes	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes
Occupation	Yes	Yes	Yes	Yes	Yes
Region effects	Yes	Yes	Yes	Yes	Yes
Observations	4,211	4,156	4,219	4,220	4,220
Pseudo R-squared	0.0713	0.0663	0.0104	0.0197	0.0542

Notes: The unit of observation is an individual. Coefficients from the first two columns are derived from a logit estimation, and the coefficients from the last three columns are derived from an ordered logit estimation. Education includes five categories, and occupation includes nine categories. Region effects include all of the regions of Yugoslavia - Bosnia, Croatia, Macedonia, Montenegro, Serbia (including Kosovo) and Slovenia. Columns 1 and 2 include a subset of answers to the question “What is the meaning of the term Yugoslavism to you?”. Answer categories are: “Only citizenship”; “Formation of a new nation” (column 1); “Supra-national identity (higher form of nationality)” (column 2); “Protest against national antagonism”; “A way to show similarity to other nationalities and ethnicities”; “Expression of patriotism, attachment to the system”; “Doesn’t hold meaning to me”; “I don’t know”. The respondents could choose only one answer to that question. Dependent variables in columns 3-5 are answers to the question “How important is your sense of belonging to these entities?”. Answer categories are: 1 = Unimportant; 2 = A little important; 3 = Average important; 4 = Quite important; 5 = Very important. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: Data come from a 1990 survey on the attitudes of Yugoslav citizens towards the economic, social and political developments in Yugoslavia published by Baćević (1991).

Yugoslavs were more likely to answer that “Yugo-Slavism” signified nation-formation to them (column 1), and a supra-national identity (column 2). Moreover, self-identified Yugoslavs were feeling less attached to their local (column 3) and regional (column 4) communities, and were feeling more attached to the wider Yugoslav community (column 5). The existing evidence thus firmly points to the conclusion that Yugoslav identification signified national sentiment and national unity.

5 Channel: Intermarriage

This section analyses a channel through which ethnic diversity had an impact on Yugoslav sentiment. As stated in section 2.2, sociologists have long perceived intermarriage as a central issue in ethnic relations. Intermarriage weakens the delineation of ethnic boundaries, and thereby decreases the salience of ethnic identities (Davis, 1991; Qian and Lichter, 2007). Furthermore, intermarried individuals, and especially their children, are less likely to identify with a single ethnicity (Waters, 1990; Xie and Goyette, 1997). Both reasons suggest that identification with the hybrid Yugoslav identity provided an alternative to forcing a single ethnic choice to intermarried individuals and their children.

If this holds, the logical step is to question why would ethnic diversity influence intermarriage. It could do so for two reasons. First, ethnic diversity determines the potential number of interethnic matches (supply). In addition, ethnic diversity might also influence preferences

Table 9: Testing the intermarriage channel, 3SLS and OLS estimates

	(1)	(2)	(3)	(4)
	3SLS			OLS
	1st stage Ethnic fractionalization	2nd stage Intermarriage	3rd stage Fraction of Yugoslavs	Fraction of Yugoslavs
Border changes, 1421-1816	0.0550** (0.026)			
Ethnic fractionalisation		0.5039*** (0.277)		0.0013 (0.007)
Intermarriage			0.3510*** (0.073)	0.4956*** (0.034)
Baseline controls	Yes	Yes	Yes	Yes
Observations	109	109	109	109
R-squared	0.293	0.598	0.896	0.915

Notes: The analysis is restricted to Bosnia-Herzegovina since the intermarriage data is not available for other regions. Intermarriage is the ratio of intermarried individuals. Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence. These controls are identical to those in column 15 of table 2. The data and the data sources are described in detail in appendix A.1. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

towards intermarriage (demand). Irrespective of the exact channel, which I cannot identify, isolating the overall mechanism of intermarriage is a challenging task due to endogeneity concerns. I begin by using the following system of three equations:

$$\begin{aligned}
e_i &= \lambda + \zeta s_i + \eta X_i + \nu_i \\
m_i &= \delta + \beta e_i + \kappa X_i + \iota_i \\
y_i &= \alpha + \phi m_i + \gamma X_i + \epsilon_i
\end{aligned} \tag{4}$$

In this system, the first stage predicts ethnic diversity by border changes. The part of the variation in ethnic diversity caused by border changes then predicts in the second stage the incidence of intermarriage, m_i (ratio of intermarried individuals to total married individuals). Finally, in the third stage, this variation in intermarriages predicts the share of Yugoslavs. I estimate this system of three equations with three-stage least squares (3SLS), following [Becker and Woessmann \(2009\)](#). This model accentuates the three-stage character of the argument that ethnic diversity influenced Yugoslav sentiment through intermarriage.

I report the 3SLS results in table 9. In column 1, border changes are positively associated with ethnic diversity. Variation in ethnic diversity caused by border changes has a positive impact on intermarriages in column 2. Finally, in column 3, the variation in intermarriages caused by ethnic diversity has a positive impact on Yugoslav identification. The coefficient is substantial and significant at 1%. It implies that moving from zero intermarriage to the sample mean of 0.07 increases the fraction of Yugoslavs by 2,5 percentage points, relative to the overall sample mean of 0.027.²⁴

Even if the exclusion restrictions hold, my specification assumes that ethnic diversity effects operated exclusively through intermarriage. This is, admittedly, problematic - it is possible that

²⁴Given the nature of the data, I cannot isolate the overall effect of intermarriage that can be attributed between the intermarried individuals and their children.

ethnic diversity operated through additional mechanisms (section 2.2). To test this possibility, I estimate an OLS model with the share of Yugoslavs as an independent variable, and ethnic diversity and intermarriage as dependent variables. The results are shown in column 4 of table 9, where the coefficient on intermarriage is large and significant. Once I control for intermarriage, ethnic diversity loses all of its association with Yugoslav sentiment - the coefficient is close to zero and insignificant. This suggests that the key mechanism linking ethnic diversity and national identity is intermarriage.

The problem with the OLS model, however, is that intermarriage is likely endogenous, just like ethnic diversity is. While I can instrument ethnic diversity with border changes, I do not have an instrument for intermarriage. I thus use a different strategy to assess the possible presence of alternative ethnic diversity effects.

I use the casual mediation analysis developed by [Dippel et al. \(2017\)](#). Within a 2SLS setting, they modify the standard mediation model by adding an instrument that causes the treatment. The advantage of this mediation approach over the standard IV method is that, with a single instrument (border changes), the effects of both the treatment (ethnic diversity) and the mediator (intermarriage) on the final outcome (Yugoslav sentiment) can be estimated. This allows me to establish causality, while unpacking the mechanism that connects ethnic diversity and Yugoslav sentiment.

In appendix A.14, I present the mediation analysis. The results imply that intermarriage explains close to 60 per cent of the overall effect of ethnic diversity. One could attribute the remaining effect to the other mechanisms documented in section 2.2. While each of the empirical methods I use can be criticized on its own, the totality of evidence suggests that intermarriage is a key channel through which ethnic diversity influenced Yugoslav identity.

6 Addendum: Yugoslav identity and the Bosnian War of 1992-1995

In this section, I explore some of the consequences of Yugoslav identity. The causes and economic consequences of civil conflict are extensively researched (see [Blattman and Miguel \(2010\)](#) for a literature survey). Ethnic divisions undoubtedly increase the incidence of civil wars ([Montalvo and Reynal-Querol, 2005b,a](#); [Esteban and Ray, 2011](#); [Esteban et al., 2012](#); [Desmet et al., 2012](#)). The corollary is that identification with a nation should reduce the likelihood and intensity of civil war. In order to examine this corollary, I analyse the correlation between Yugoslav identity and the intensity of conflict during the Bosnian War of 1992-1995.

Table 10 present the results. The dependent variable is the ratio of war casualties to the population. The controls follow the specification of [Montalvo and Reynal-Querol \(2005b\)](#), and they are reported in the table. The first column shows that municipalities characterised by a larger fraction of self- declared Yugoslavs were characterised by a lower intensity of conflict. The effect is significant at 1%. Even though this relationship is non-casual, it seems plausible that individuals that identified with the wider nation were less likely to be an object and subject of an ethnic-based conflict, perhaps by fleeing the country.

To control for ethnic divisions, in column 2, I include the ethnic fractionalisation index.

Table 10: OLS estimates, Yugoslav identity and civil war in Bosnia-Herzegovina (1992-1995), dependent variable: population share of war casualties

	(1)	(2)	(3)	(4)
Fraction of Yugoslavs	-0.2087*** (0.072)	-0.3437*** (0.079)	-0.2873*** (0.073)	-0.1357** (0.061)
Ethnic fractionalisation		0.0497*** (0.010)		-0.0985*** (0.037)
Ethnic polarisation			0.0390*** (0.007)	0.0965*** (0.026)
Controls	Yes	Yes	Yes	Yes
Observations	109	109	109	109
R-squared	0.064	0.155	0.211	0.248

Notes: Output per capita is included as a proxy for the opportunity cost of engaging in conflict. Share of working-age population is included as a proxy for recruitable soldiers. Expenditure on public goods p.c. is included as a proxy for the benefits of conflict, since it indicates possible revenue extraction. Terrain roughness is included as a proxy for potential safe heaven for military factions in hostile territory. The fraction of labour employed in labour-managed firms is included as a proxy for local (work-place) democracy. Finally, I control for the Yugoslav National Army presence, which was “hijacked” by the Serbian military forces. The data and the data sources are described in detail in appendix A.1. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The coefficient on the fraction of Yugoslavs remains negative, while the coefficient on ethnic fractionalisation is positive. In column 3, I substitute the ethnic polarization index for the ethnic fractionalisation index. The results are similar to those in column 2. Finally, in column 4, I include jointly the ethnic fractionalisation and polarisation indexes. The coefficient on the fraction of Yugoslavs remains negative. However, the coefficient on ethnic fractionalisation now turns negative, while the coefficient on ethnic polarisation remains positive. Thus, even when conditioned by ethnic fractionalisation and ethnic polarisation, Yugoslav identification remains negatively correlated with the intensity of conflict. These estimates cannot be considered causal, and the size of the coefficients are likely biased. Nevertheless, the direction of the coefficients seems plausible - they mirror those of [Montalvo and Reynal-Querol \(2005b,a\)](#).

7 Conclusion

This paper isolates the impact of ethnic diversity on national identity in the historical setting of the multi-ethnic Yugoslavia. I find that ethnically diverse municipalities were conducive towards Yugoslav self-identification because they stimulated ethnic intermarriage. I interpret this finding to mean that the more hybrid Yugoslav identity provided an alternative to forcing a single choice to people with conflicting pressures on their identity.

Did the formation of Yugoslav identity contain wider consequences? Aligned with the notion that nation-building can reduce the incidence of ethnic conflict, I find that municipalities that

were characterised by a larger number of self-declared Yugoslavs experienced a lower intensity of conflict during the Bosnian War of 1992-1995. Ultimately, there were not enough Yugoslavs to preserve the country, and avoid the civil war.

Ethnic diversity matters. However, nation-building policies matter as well. Such policies, however, were not attempted in Yugoslavia. The overwhelming fraction of population could not identify with a Yugoslav nation, and the country disintegrated.

It is natural to wonder whether the findings of this paper contain wider significance and whether they can inform policy-making. The results suggest that nation building strategies in ethnically divided societies could be particularly effective in ethnically diverse areas, because such areas tend to be more susceptible to the idea of a common nation. Moreover, to mitigate the risk of ethnic tension and conflict, the empirical results of this paper imply a role for a specific policy. Such policy could be the promotion of ethnic intermarriage, as the surest path to a world without ethnic hatred and conflict is a world without ethnicities.

References

- AHLERUP, P. AND O. OLSSON (2012): “The Roots of Ethnic Diversity,” *Journal of Economic Growth*, 17, 71–102.
- AKERLOF, G. A. AND R. E. KRANTON (2000): “Economics and Identity,” *Quarterly Journal of Economics*, 115, 715–753.
- (2002): “Identity and Schooling: Some Lessons for the Economics of Education,” *Journal of Economic Literature*, 40, 1167–1201.
- (2005): “Identity and the Economics of Organizations,” *Journal of Economic Perspectives*, 19, 9–32.
- ALESINA, A., A. DEVLEESCHAUWER, W. EASTERLY, S. KURLAT, AND R. WACZIARG (2003): “Fractionalization,” *Journal of Economic Growth*, 8, 155–194.
- ALESINA, A. AND E. LA FERRARA (2005): “Ethnic Diversity and Economic Performance,” *Journal of Economic Literature*, 43, 762–800.
- ALESINA, A. F., B. REICH, AND A. RIBONI (2017): “Nation-Building, Nationalism and Wars,” NBER Working Paper Series No. 23435.
- ALI, M., O.-H. FJELDSTAD, B. JIANG, AND A. B. SHIFA (2018): “Colonial Legacy, Statebuilding and the Salience of Ethnicity in Sub-Saharan Africa,” *Economic Journal*, Advance online publication, doi.org/10.1111/econj.12595.
- ALLPORT, G. (1954): *The Nature of Prejudice*, Cambridge, MA: Perseus Books.
- ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools,” *Journal of Political Economy*, 113, 151–184.
- ANDERSON, B. (1983): *Imagined Communities: Reflections on the Origin and Spread of Nationalism*, London: Verso.
- ASHRAF, Q. AND O. GALOR (2013): “Genetic Diversity and the Origins of Cultural Fragmentation,” *American Economic Review: Papers and Proceedings*, 103, 528–533.
- BAĆEVIĆ, L. (1991): *Jugoslavija na Kriznoj Prekretnici*, Beograd: Universitet u Beogradu, Institut Društvenih Nauka.
- BALL, P., E. TABEAU, AND P. VERWIMP (2007): “The Bosnian Book of Dead: Assessment of the Database (Full Report),” Households in Conflict Network, University of Sussex, The Institute of Development Studies.
- BANAC, I. (1984): *The National Question in Yugoslavia: Origins, History, Politics*, Ithaca, NY: Cornell University Press.
- BANDIERA, O., M. MOHNEN, I. RASUL, AND M. VIARENGO (2019): “Nation-Building Through Compulsory Schooling During the Age of Mass Migration,” *Economic Journal*, 129, 62–109.
- BECKER, S. O., K. BOECKH, C. HAINZ, AND L. WOESSMANN (2016): “The Empire is Dead, Long Live the Empire! Long-run Persistence of Trust and Corruption in the Bureaucracy,” *Economic Journal*, 126, 40–74.
- BECKER, S. O. AND L. PASCALI (2019): “Religion, Division of Labor, and Conflict: Anti-semitism in Germany over 600 Years,” *American Economic Review*, 109, 1764–1804.

- BECKER, S. O. AND L. WOESSMANN (2009): “Was Weber Wrong? A Human Capital Theory of Protestant Economic History,” *Quarterly Journal of Economics*, 124, 531–596.
- BELLOWS, J. AND E. MIGUEL (2009): “War and Local Collective Action in Sierra Leone,” *Journal of Public Economics*, 93, 1144–1157.
- BISIN, A. AND T. VERDIER (2001): “The Economics of Cultural Transmission and the Dynamics of Preferences,” *Journal of Economic Theory*, 97, 298–319.
- BLATTMAN, C. AND E. MIGUEL (2010): “Civil War,” *Journal of Economic Literature*, 48, 3–57.
- BLOUIN, A. AND S. W. MUKAND (2019): “Erasing Ethnicity? Nation Building, (Mis)Trust and the Salience of Identity in Rwanda,” *Journal of Political Economy*, 127, 1008–1062.
- BOISJOLY, J., G. J. DUNCAN, M. K. D. M. LEVY, AND J. ECCLES (2006): “Empathy or Antipathy? The Impact of Diversity,” *American Economic Review*, 96, 1890–1905.
- BOUGAREL, X. (2003): “Bosnian Muslims and the Yugoslav Idea,” in *Yugoslavism: Histories of a Failed Idea*, ed. by D. Djokic, London: Hurst Company, 100–114.
- BURG, S. L. (1983): *Conflict and Cohesion in Socialist Yugoslavia*, Princeton, NJ: Princeton University Press.
- BURG, S. L. AND M. L. BERBAUM (1989): “Community, Integration, and Stability in Multinational Yugoslavia,” *American Political Science Review*, 83, 535–554.
- CASELLI, F. AND W. J. COLEMAN (2013): “On the Theory of Ethnic Conflict,” *Journal of the European Economic Association*, 11, 161–192.
- CENTRALNI KOMITET SAVEZA KOMUNISTA HRVATSKE (1962): “Komisija za Organizaciona Pitanja - Statistika, Broj OK SK; Trenutno Stanje i Prijedlozi za Nove Službenike u Komitetima,” Internal Documents, croatian State Archives, HR-HAD-1220.2.3.1.1.7, box 11.
- CINNIRELLA, F. AND R. SCHUELER (2018): “Nation Building: The Role of Central Spending in Education,” *Explorations in Economic History*, 67, 18–39.
- CLOTS-FIGUERAS, I. AND P. MASELLA (2013): “Education, Language, and Identity,” *Economic Journal*, 123, 332–357.
- COLLEY, L. (1992): *Britons: Forging the Nation, 1707-1837*, New Haven, CT: Yale University Press.
- CONLEY, T. G., C. B. HANSEN, AND P. E. ROSSI (2012): “Plausibly Exogenous,” *Review of Economics and Statistics*, 94, 260–272.
- CONNOR, W. (1984): *The National Question in Marxist-Leninist Theory and Strategy*, Princeton, NJ: Princeton University Press.
- DAVIS, F. (1991): *Who Is Black?: One Nation’s Definition*, University Park, PA : Pennsylvania State University Press.
- DELL, M. AND P. QUERUBIN (2018): “Nation Building Through Foreign Intervention: Evidence from Discontinuities in Military Strategies,” *Quarterly Journal of Economics*, 133, 701–764.
- DESMET, K., I. ORTUNO-ORTIN, AND R. WACZIARG (2012): “The Political Economy of Linguistic Cleavages,” *Journal of Development Economics*, 97, 583–597.
- DIMITRIJEVIĆ, B. (2001): “Armija i Jugoslovenski Identitet, 1945-1992.” *Vojno Delo*, 53, 141–154.

- DIPPEL, C., R. GOLD, S. HEBLICH, AND R. PINTO (2017): “Instrumental Variables and Causal Mechanisms: Unpacking the Effect of Trade on Workers and Voters,” NBER Working Paper Series No. 23209.
- DJOKIĆ, D. (2003): “Introduction,” in *Yugoslavism: Histories of a Failed Idea*, ed. by D. Djokic, London: Hurst & Company, chap. 1.
- DRAGONER (Online): “Jugoslavenska Narodna Armija (JNA) - Yugoslav People’s Army, OOB 1990,” <http://elektron.tmf.bg.ac.rs/bojan/JNA%20OOB.pdf>, accessed: 3rd of August 2016.
- DRŽAVNI ZAVOD ZA STATISTIKU (1992a): *Popis Stanovništva 1991: Stanovništvo Prema Narodnosti po Naseljima*, Zagreb: Državni Zavod za Statistiku.
- (1992b): *Popis Stanovništva 1991: Stanovništvo Prema Školskoj Spremi, Pismenosti i Spolu po Naseljima*, Zagreb: Državni Zavod za Statistiku.
- (1994a): *Popis Stanovništva 1991: Aktivno Stanovništvo u Zemlji koje Obavlja Zanimanje, Prema Području Djelatnosti, po Naseljima*, Zagreb: Državni Zavod za Statistiku.
- (1994b): *Popis Stanovništva 1991: Stanovništvo Prema Spolu i Starosti, po Naseljima*, Zagreb: Državni Zavod za Statistiku.
- DUGAČKI, V. (2009): “Historiografija o Česima u Hrvatskoj,” *Historijski Zbornik*, 61, 235–252.
- ESTEBAN, J., L. MAYORAL, AND D. RAY (2012): “Ethnicity and Conflict: An Empirical Study,” *American Economic Review*, 102, 1310–1342.
- ESTEBAN, J. AND D. RAY (2011): “Linking Conflict to Inequality and Polarization,” *American Economic Review*, 101, 1345–1374.
- EUROPEAN ENVIRONMENT AGENCY (Online): “Elevation Map of Europe,” <https://www.eea.europa.eu/data-and-maps>, accessed: 28h of March 2018.
- FEDERALNI ZAVOD ZA STATISTIKU (1993): *Popis Stanovništva 1991: Etnička Obilježja Stanovništva, Rezultati za Republiku i po Opštinama*, Sarajevo: Federalni Zavod za Statistiku.
- (1994): *Popis Stanovništva 1991: Porodice u Republici Bosni i Hercegovini, Rezultati za Republiku i po Opštinama/Općinama*, Sarajevo: Federalni Zavod za Statistiku.
- (1998a): *Popis Stanovništva 1991: Nacionalni Sastav Stanovništva po Naseljenim Mjestima*, Sarajevo: Federalni Zavod za Statistiku.
- (1998b): *Popis Stanovništva 1991: Stanovništva Prema Starosti i Spolu po Naseljenim Mjestima*, Sarajevo: Federalni Zavod za Statistiku.
- (1998c): *Popis Stanovništva 1991: Uporedni Podaci 1971, 1981, 1991*, Sarajevo: Federalni Zavod za Statistiku.
- FERRARA, A. (2018): “World War II and African American Socioeconomic Progress,” CAGE Working Paper Series No. 387.
- FOOD AND AGRICULTURE ORGANIZATION (Online): “Agricultural Suitability and Potential Yields,” <http://www.fao.org/nr/gaez/about-data-portal/agricultural-suitability-and-potential-yields/en/>, accessed: 25th of March 2016.
- GELLNER, E. (1983): *Nations and Nationalism*, Ithaca, NY: Cornell University Press.
- GIBBONS, S. AND H. G. OVERMAN (2012): “Mostly Pointless Spatial Econometrics?” *Journal of Regional Science*, 52, 172–191.

- GOATI, V. (1998): *Elections to the Federal and Republican Parliaments of Yugoslavia (Serbia and Montenegro) 1990-1996. Analyses, Documents and Data*, Berlin: Ed. Sigma.
- GROSFELD, I. AND E. ZHURAVSKAYA (2015): “Cultural vs. Economic Legacies of Empires: Evidence from the Partition of Poland,” *Journal of Comparative Economics*, 43, 55–75.
- GROSJEAN, P. (2011): “The Weight of History on European Cultural Integration: A Gravity Approach,” *American Economic Review: Papers and Proceedings*, 101, 504–508.
- HOBBSAWM, E. (1991): *Nations and Nationalism Since 1780: Programme, Myth, Reality*, Cambridge: Cambridge University Press.
- HODSON, R., G. MASSEY, AND D. SEKULIĆ (1994a): “National Tolerance in the Former Yugoslavia,” *American Journal of Sociology*, 99, 1534–1558.
- (1994b): “Who Were the Yugoslavs? Failed Sources of a Common Identity in the Former Yugoslavia,” *American Sociological Review*, 59, 83–97.
- HÖTTE, H. (2015): *Atlas of Southeast Europe: Geopolitics and History. Volume One: 1521–1699*, Leiden: Brill.
- INDEX.HR (2008, December 8): “Lepa Brena: Nisam ni Hrvatica ni Srпкиnja, ja sam Jugoslavenka!” <http://www.index.hr/xmag/clanak/lepa-brena-nisam-ni-hrvatica-ni-srпкиnja-ja-sam-jugoslavenka/412754.aspx>, accessed: 22nd of May 2017.
- JHA, S. (2013): “Trade, Institutions, and Ethnic Tolerance: Evidence from South Asia,” *American Political Science Review*, 107, 806–832.
- JOKIĆ, G. (1986): *Jugoslavija: Spomenici Revolucije, Turistički Vodič*, Beograd: Turistička Štampa.
- JOVIĆ, D. (2009): *Yugoslavia: A State that Withered Away*, West Lafayette, IN: Purdue University Press.
- JURAJDA, S. AND D. KOVAČ (2016): “What’s in a Name in a War,” IZA Discussion Paper No. 10331.
- JURKOVIĆ, I. (2005): “Šesnaestostoljetna Hrvatska Raseljenička Kriza i Moderna Sociološka Terminologija,” *Društvena Istraživanja*, 14, 183–199.
- KARDELJ, E. (1979): *The Nations and Socialism*, Beograd: STP.
- KEEGAN, J. (1993): *A History of Warfare*, London: Hutchinson.
- KORDIĆ, S. (2010): *Jezik i Nacionalizam*, Zagreb: Duriex.
- KOREN, S. (2012): *Politika Povijesti u Jugoslaviji (1945-1960): Komunistička Partija Jugoslavije, Nastava Povijesti, Historiografija*, Zagreb: Srednja Europa.
- KOYAMA, M., R. JEBWAB, AND N. D. JOHNSON (2019): “Negative Shocks and Mass Persecutions: Evidence from the Black Death,” *Journal of Economic Growth*, Advance online publication, doi.org/10.1007/s10887-019-09167-1.
- LAMPE, J. R. (2000): *Yugoslavia as History, Twice there was a Country*, Cambridge: Cambridge University Press.
- LANE, F. C. (1973): *Venice, a Maritime Republic*, Baltimore, MD: Johns Hopkins University Press.

- MANNING, A. AND S. ROY (2010): "Culture Clash or Culture Club? National Identity in Britain," *Economic Journal*, 120, F72–F100.
- MARKOWITZ, F. (2007): "Census and Sensibilities in Sarajevo," *Comparative Studies in Society and History*, 49, 40–73.
- MICHALOPOULOS, S. (2012): "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102, 1508–1539.
- MILISAVAC, Z. (1971): *Jugoslovenski Književni Leksikon*, Novi Sad: Matic Srpska.
- MONTALVO, J. G. AND M. REYNAL-QUEROL (2005a): "Ethnic Diversity and Economic Development," *Journal of Development Economics*, 76, 293–323.
- (2005b): "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, 95, 796–816.
- (2017): "Ethnic Diversity and Growth: Revisiting the Evidence," Universitat Pompeu Fabra-ICREA, Economics Working Paper Series, No. 1585.
- MURPHEY, R. (1996): *Ottoman Warfare: 1500-1700*, London: UCL Press.
- NOVAKOVIĆ, R. (1965): *Školski Istorijski Atlas*, Beograd: Zavod za Izdavanje Udžbenika SR Srbije.
- PAPKE, L. E. AND J. M. WOOLDRIDGE (1996): "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates," *Journal of Applied Econometrics*, 11, 619–632.
- PAVLOVIĆ, M. (2015): "The Migration of Population from the Territory of Serbia to Turkey in a Historical Perspective," *Bulletin of the Institute of Ethnography SASA*, 63, 581–579.
- PERKOVIĆ, A. (2011): *Sedma Republika: Pop Kultura u YU Raspadu*, Zagreb: Novi Liber.
- POPOVIĆ, D. (2010): "Youth Labor Action (*Omladinska Radna Ackija, ORA*) as Ideological Holiday-Making," in *Yugoslavia's Sunny Side: A History of Tourism in Socialism (1950s-1980s)*, ed. by H. Grandits and K. Taylor, Budapest: Central European University, 279–302.
- QIAN, Z. AND D. T. LICHTER (2007): "Social Boundaries and Marital Assimilation: Interpreting Trends in Racial and Ethnic Intermarriage," *American Sociological Review*, 72, 68–94.
- RAMET, S. P. (1992): *Balkan Babel: Politics, Culture, and Religion in Yugoslavia*, Boulder, CO: Westview Press.
- REGAN, K. AND T. KANIŠKI (2003): *Hrvatski Povijesni Atlas*, Zagreb: Leksikografski Zavod Miroslav Krleža.
- REPUBLIČKI ZAVOD ZA STATISTIKU (2003): *Popis Stanovništva 2002: Školska Sprema i Pismenost, Prema Opštinama*, Beograd: Republički Zavod za Statistiku.
- ROTHENBERG, G. E. (1966): *The Military Border in Croatia, 1740-1881: A Study of an Imperial Institution*, Chicago, IL: University of Chicago Press.
- RUTHNER, C. (2019): "Habsburg's Only Colony? Bosnia-Herzegovina and Austria-Hungary, 1878-1918," *SEEU Review*, 13, 2–14.
- SAVEZNI ZAVOD ZA STATISTIKU (1964): *Statistički Godišnjak SFR Jugoslavije 1964*, Beograd: Savezni Zavod za Statistiku.

- (1966a): *Statistički Bilten Broj 418.: Omladinske Radne Akcije 1964. i 1965*, Beograd: Savezni Zavod za Statistiku.
- (1966b): *Statistički Godišnjak SFR Jugoslavije 1965*, Beograd: Savezni Zavod za Statistiku.
- (1966c): *Statistički Godišnjak SFR Jugoslavije 1966*, Beograd: Savezni Zavod za Statistiku.
- (1966d): *Žrtve Rata, 1941-1945 (Rezultati Popisa)*, Beograd: Savezni Zavod za Statistiku.
- (1974): *Popis Stanovništva 1971: Etnička, Prosvetna i Ekonomska Obilježlja Stanovništva i Domaćinstva Prema Broju Članova, Rezultati po Opštinama*, Beograd: Savezni Zavod za Statistiku.
- (1982): *Statistički Bilten Broj 1411: Borci, Vojni Invalidi i Porodice Palih Boraca - Korisnici Osnovnih Prava po Saveznim Propisima, 31.12.1981.*, Beograd: Savezni Zavod za Statistiku.
- (1991a): *Popis Stanovništva 1991: Stanovništvo Prema Veroispovesti*, Beograd: Savezni Zavod za Statistiku.
- (1991b): *Statistički Godišnjak SFR Jugoslavije 1991*, Beograd: Savezni Zavod za Statistiku.
- (1993a): *Popis Stanovništva 1991: Stanovništvo Prema Nacionalnoj Pripadnosti*, Beograd: Savezni Zavod za Statistiku.
- (1993b): *Popis Stanovništva 1991: vol. 24*, Beograd: Savezni Zavod za Statistiku.
- SEKULIĆ, D. (2004): "Civic and Ethnic Identity: The Case of Croatia," *Ethnic and Racial Studies*, 27, 455–483.
- SIBER, I. (1997): *The 1990 and 1992/93 Sabor Elections in Croatia: Analyses, Documents and Data. Series Founding Elections in Eastern Europe*, Berlin: Ed. Sigma.
- SILVA, J. M. S. AND S. TENREYRO (2006): "The Log of Gravity," *Review of Economics and Statistics*, 88, 641–658.
- STATISTIČNI URAD REPUBLIKE SLOVENIJE (1994a): *Popis Prebivalstva 1991: Aktivno Prebivalstvo, ki Pravlja Poklic, po Dejavnosti (brez Oseb, ki Delajo v Tujini in Družinskih Članov, ki z Njimi Živijo v Tujini)*, Ljubljana: Statistični Urad Republike Slovenije.
- (1994b): *Popis Prebivalstva 1991: Prebivalstvo po Narodnostni Pripadnosti*, Ljubljana: Statistični Urad Republike Slovenije.
- (1994c): *Popis Prebivalstva 1991: Prebivalstvo po Starosti in Spolu*, Ljubljana: Statistični Urad Republike Slovenije.
- (1994d): *Popis Prebivalstva 1991: Prebivalstvo po Veroizpovedi*, Ljubljana: Statistični Urad Republike Slovenije.
- (1994e): *Popis Prebivalstva 1991: Prebivalstvo, Staro 15 Let ali več, po Spolu, Izobrazbi in Pismenosti*, Ljubljana: Statistični Urad Republike Slovenije.
- STEWART-STEINBERG, S. (2007): *The Pinocchio Effect: On Making Italians, 1860-1920*, Chicago, IL: Chicago University Press.

- SUBTELNY, O. (2007): "The Habsburg and Russian Empires: Some Comparisons and Contrasts," in *Empire and Society: New Approaches to Russian History*, ed. by H. Teruyuki and K. Matsuzato, Sapporo: Slavic Research Center, Hokkaido University, 73–92.
- TOKAČA, M. (2012): *Bosanska Knjiga Mrtvih: Ljudski Gubici u Bosni i Hercegovini, 1991-1995*, Sarajevo: Istraživačko Dokumentacioni Centar.
- VALENTIĆ, M. (1990): "Turski Ratovi i Hrvatska Dijaspore u XVI. Stoljecu," *Senjski Zbornik*, 17, 45–60.
- VOIGTLÄNDER, N. AND H.-J. VOTH (2012): "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany," *Quarterly Journal of Economics*, 127, 1339–1392.
- ŠEHIĆ, Z. AND I. TEPIĆ (2002): *Povijesni Atlas Bosne i Hercegovine*, Sarajevo: Sejtarija.
- ŽERJAVIĆ, V. (1993): "Doseljavanja i Iseljavanja s Područja Istre, Rijeke i Zadra u Razdoblju 1910-1971," *Društvena Istraživanja*, 2, 631–656.
- WACHTEL, A. (1998): *Making a Nation, Breaking a Nation: Literature and Cultural Politics in Yugoslavia*, Stanford, CA: Stanford University Press.
- WATERS, M. C. (1990): *Ethnic Options: Choosing Identities in America*, Berkeley, CA: University of California Press.
- WEBER, E. (1976): *Peasants into Frenchmen*, Stanford, CA: Stanford University Press.
- XIE, Y. AND K. GOYETTE (1997): "The Racial Identification of Biracial Children with One Asian Parent: Evidence from the 1990 Census," *Social Forces*, 76, 547–570.
- ZAVOD ZA STATISTIKU (2005): *Popis Stanovništva 2003: Školska Sprema i Pismenost, Rezultati po Opštinama*, Podgorica: Zavod za Statistiku.

A Appendix

A.1 Data

This section documents the construction of each variable that I use in this paper, alongside the data sources.

Population and ethnicity - Population and ethnicity data for municipalities is taken from the 1991 population census books of the successor states of Yugoslavia. For Bosnia-Herzegovina: *Popis Stanovništva 1991: Nacionalni Sastav Stanovništva po Naseljenim Mjestima* (Federalni Zavod za Statistiku, 1998a). For Croatia: *Popis Stanovništva 1991: Stanovništvo Prema Narodnosti po Naseljima* (Državni Zavod za Statistiku, 1992a). For Montenegro and Serbia: *Popis Stanovništva 1991: Stanovništvo Prema Nacionalnoj Pripadnosti* (Savezni Zavod za Statistiku, 1993a). For Slovenia: *Popis Prebivalstva 1991: Prebivalstvo po Narodnosti Pripadnosti* (Statistični Urad Republike Slovenije, 1994b).

Border changes - The baseline instrument, border changes, is defined as the amount of political entities a municipality had belonged to between 1421 and 1816. For Bosnia-Herzegovina, Croatia, Slovenia, and most of Montenegro, I derive the instrument from a historical atlas which focuses on Croatia, but also covers surrounding regions - *Povijesni Atlas Hrvatske* (Historical Atlas of Croatia, (Regan and Kaniški, 2003)). For Serbia, and partly for Montenegro, I derive the instrument from a set of additional historical atlases (Novaković, 1965; Šehić and Tepić, 2002; Hötte, 2015). I have included all border changes that are reported in these atlases. The baseline instrument, however, excludes reconquests of territories to minimise measurement issues. Reconquests are under-reported in historical atlases, particularly if ruler managed to hold to a territory for a brief period of time in the presence of military offensives and counter-offensives (Regan and Kaniški, 2003).

Religiosity and Muslims - Data on religious structure of municipalities is taken from the 1991 population census books of the successor states of Yugoslavia. For Bosnia-Herzegovina: *Popis Stanovništva 1991: Etnička Obilježlja Stanovništva, Rezultati za Republiku i po Opštinama* (Federalni Zavod za Statistiku, 1993). For Croatia: *Popis Stanovništva 1991: Stanovništvo Prema Vjeroispovijesti i Materinskom Jeziku, po Naseljima* (Državni Zavod za Statistiku, 1992a). For Montenegro and Serbia: *Popis Stanovništva 1991: Stanovništvo Prema Veroispovesti* (Savezni Zavod za Statistiku, 1991a). For Slovenia: *Popis Prebivalstva 1991: Prebivalstvo po Veroizpovedi* (Statistični Urad Republike Slovenije, 1994d).

Language - Data on the self-reported mother tongue is taken from the 1991 population census books of the successor states of Yugoslavia. For Bosnia-Herzegovina: *Popis Stanovništva 1991: Etnička Obilježlja Stanovništva, Rezultati za Republiku i po Opštinama* (Federalni Zavod za Statistiku, 1993). For Croatia: *Popis Stanovništva 1991: Stanovništvo Prema Vjeroispovijesti i Materinskom Jeziku, po Naseljima* (Državni Zavod za Statistiku, 1992a). For Slovenia: *Popis Prebivalstva 1991: Prebivalstvo po Veroizpovedi* (Statistični Urad Republike Slovenije, 1994d).

Output p.c. - Municipality output data is taken from the 1991 Yugoslav Statistical Yearbook: *Statistički Godišnjak SFR Jugoslavije 1991* (Savezni Zavod za Statistiku, 1991b). Output data is reported for 1989. The successor states of Yugoslavia had stopped estimating output at municipal level afterwards. Thus, the final year for which output has been estimated is the year 1989. As in other socialist countries, Yugoslav statisticians excluded some services from the value of output - education, healthcare, and housing. Output is normalised by the 1991 population.

Population density - Population density of municipalities is constructed by dividing population by the administrative area (km^2). Data on administrative areas is taken from the *Statistički Godišnjak SFR Jugoslavije 1991* (Savezni Zavod za Statistiku, 1991b).

Average years of schooling - Educational attainment for each municipality is estimated for the population aged 15 and above. With the exception of Montenegro and Serbia, average years of schooling for the population of each municipality is constructed from the 1991 population

census books of the successor states of Yugoslavia. For Bosnia-Herzegovina: *Popis Stanovništva 1991: Uporedni Podaci 1971, 1981, 1991* (Federalni Zavod za Statistiku, 1998c). For Croatia: *Popis Stanovništva 1991: Stanovništvo Prema Školskoj Spremi, Pismenosti i Spolu po Naseljima* (Državni Zavod za Statistiku, 1992b). For Slovenia: *Popis Prebivalstva 1991: Prebivalstvo, Staro 15 Let ali vec, po Spolu, Izobrazbi in Pismenosti* (Statistični Urad Republike Slovenije, 1994e). For Serbia and Montenegro, the 1991 population census does not record educational attainment at municipal level. As such, I have interpolated educational attainment for each municipality from the population census from the two nearest population censuses that reported this information. For Montenegro, I have interpolated the 1991 values from *Popis Stanovništva 1971: Etnička, Prosvetna i Ekonomska Obilježja Stanovništva i Domaćinstva Prema Broju Članova, Rezultati po Opštinama* (Savezni Zavod za Statistiku, 1974) and *Popis Stanovništva 2003: Školska Sprema i Pismenost, Rezultati po Opštinama* (Zavod za Statistiku, 2005). For Serbia, I have interpolated the 1991 values from *Popis Stanovništva 1971: Etnička, Prosvetna i Ekonomska Obilježja Stanovništva i Domaćinstva Prema Broju Članova, Rezultati po Opštinama* (Savezni Zavod za Statistiku, 1974) and *Popis Stanovništva 2002: Školska Sprema i Pismenost, Prema Opštinama* (Republički Zavod za Statistiku, 2003).

Social sector labour - Municipality data for social sector employment as a fraction of total labour is taken from the 1991 population census books of the successor states of Yugoslavia. For Bosnia-Herzegovina: *Popis Stanovništva 1991: Uporedni Podaci 1971, 1981, 1991* (Federalni Zavod za Statistiku, 1998c). For Croatia: *Popis Stanovništva 1991: Aktivno Stanovništvo u Zemlji koje Obavlja Zanimanje, Prema Području Djelatnosti, po Naseljima* (Državni Zavod za Statistiku, 1994a). For Montenegro and Serbia *Popis Stanovništva 1991: vol. 24* (Savezni Zavod za Statistiku, 1993b). For Slovenia: *Popis Prebivalstva 1991: Aktivno Prebivalstvo, ki Pravlja Poklic, po Dejavnosti (brez Oseb, ki Delajo v Tujini in Družinskih Članov, ki z Njimi Živijo v Tujini)* (Statistični Urad Republike Slovenije, 1994a).

Youth labour actions - Data on number of individuals involved in youth labour actions is taken from *Statistički Bilten Broj 418.: Omladinske Radne Akcije 1964. i 1965* (Savezni Zavod za Statistiku, 1966a). The final year the Yugoslav statistical office has published this data for is 1965. I thus use data on individuals involved in youth labour actions in 1965, and normalise it by 1965 population taken from *Statistički Godišnjak SFR Jugoslavije 1966* (Savezni Zavod za Statistiku, 1966c).

WWII partisan veterans - Data on partisan veterans of WWII resident in a municipality is taken from *Statistički Bilten Broj 1411: Borci, Vojni Invalidi i Porodice Palih Boraca - Korisnici Osnovnih Prava po Saveznim Propisima, 31.12.1981.* (Savezni Zavod za Statistiku, 1982). This data is reported for 1981, and has not been reported again since. I thus use the 1981 data on partisan veterans resident in a municipality, and normalise it by the 1981 population taken from *Statistički Godišnjak SFR Jugoslavije 1991* (Savezni Zavod za Statistiku, 1991b).

Monuments to anti-fascism - This is a dummy variable reporting whether a municipality contained a site or event that was considered officially to be of high importance to the anti-fascist, communist-led, struggle during WWII. These were often commemorated by a memorial complex, museum, or a monument, whereby the largest memorial centres were visited by as much as 5 million people on an annual basis (Jokić, 1986). I take this data from a detailed tourist guide (446 pages) dedicated to such sites (Jokić, 1986).

Fascist terror - This is official data on number of individuals exposed to fascist terror during WWII at current municipality of residence, and that survived. It is taken from the census that collected this data - *Žrtve Rata, 1941-1945 (Rezultati Popisa)* (Savezni Zavod za Statistiku, 1966d). The survivors include those that were in internment and deportation camps, jails, war captivity, and experienced forced labour. The census was conducted in 1964, but remained a top-secret publication used for internal use by top-ranking Communist Party members only until the dissolution of Yugoslavia, as census data on individuals killed by fascists did not conform to official public figures, which were magnified. As the data on individuals exposed to fascist terror

is reported for 1964, I normalise it by the 1964 population taken from *Statistički Godišnjak SFR Jugoslavije 1965* ([Savezni Zavod za Statistiku, 1966b](#))

Public goods exp. p.c. - Expenditure on public goods includes education, healthcare and social care. Municipality data is taken from the *Statistički Godišnjak SFR Jugoslavije 1991* ([Savezni Zavod za Statistiku, 1991b](#)) for 1989 (the last reported year). Expenditure on public goods is normalised by the 1991 population.

Federal Aid - This is a dummy variable capturing whether a municipality was a recipient of federal aid. I can determine whether a municipality was a recipient aid only until 1963. Afterwards, the statistical yearbooks, or other official publications, stopped reporting this information. The data on the 1963 recipients of federal aid is taken from the *Statistički Godišnjak SFR Jugoslavije 1964* ([Savezni Zavod za Statistiku, 1964](#)).

1980's generation - I define the early 1980s youth as the fraction of population in 1991 that was between 25 and 39 years. The data for municipalities is taken from the 1991 population census books of the successor states of Yugoslavia. For Bosnia-Herzegovina: *Popis Stanovništva 1991: Stanovništvo Prema Starosti i Spolu po Naseljenim Mjestima* ([Federalni Zavod za Statistiku, 1998b](#)). For Croatia: *Popis Stanovništva 1991: Stanovništvo Prema Spolu i Starosti, po Naseljima* ([Državni Zavod za Statistiku, 1994b](#)). For Slovenia: *Popis Prebivalstva 1991: Prebivalstvo po Starosti in Spolu* ([Statistični Urad Republike Slovenije, 1994c](#)).

Army presence - *Yugoslav National Army* presence in municipalities is measured for 1990. It includes the amount of battalions, divisions, army headquarters, naval bases, airports, academies, training centres, and research facilities. I take the data from [Dragoner](#) (online).

Communist Party members - I have gathered data on Communist Party members per municipality from the internal documentation of the [Centralni Komitet Saveza Komunističke Hrvatske](#) (Central Committee of the League of Communists of Croatia). The documentation can be accessed in the Croatian State Archives. The data comes from the membership reports that municipal branches of the Party were sending to their headquarters in Zagreb. The population fraction of Party's members is reported for 1962. I have not been able to access or find data for any later year. Municipality membership reports have been removed from the documentation of the later years. The amount of Party members is normalised by the 1961 population. Data on 1961 population at the level of municipalities is taken from the same internal documentation.

Duration of Habsburg rule - Data on the duration of the historic affiliation of a municipality with the Habsburgs is derived from [Regan and Kaniški \(2003\)](#); [Novaković \(1965\)](#); [Šehić and Tepić \(2002\)](#); [Hötte \(2015\)](#).

Agricultural (wheat) suitability - Data on the land suitability for wheat production is taken from the [Food and Agriculture Organization](#) (FAO) of the United Nations. FAO provides a set of raster data covering the agroecological environment at the global level. The data provide a sufficiently high resolution to examine the average quality of land in the municipalities of Yugoslavia.

Terrain roughness - This variable is calculated as the mean difference between a central pixel and its surrounding cells. I used the European elevation (raster) data provided by [European Environment Agency](#) (EEA). The elevation data is of very high resolution (1 km x 1 km), allowing me to calculate terrain roughness across Yugoslavia.

Intermarriage - Municipal data on the stock ratio of ethnic intermarriages to total marriages is taken from the 1991 population census of Bosnia-Herzegovina (*Popis Stanovništva 1991: Porodice u Republici Bosni i Hercegovini, Rezultati za Republiku i po Opštinama/Općinama*, [Federalni Zavod za Statistiku \(1994\)](#)). Unfortunately, other Yugoslav republics did not gather data on the stock of intermarriages at municipal level. The fraction of intermarried couples is multiplied by two in order to derive a fraction of intermarried people.

War Casualties - This variable is the amount of people that died in the area of a municipality during the Bosnian Civil War of 1992-1995. It is normalised by the 1991 population. These war casualties did not necessarily hold permanent residence at the location of their death. In

2012, the nongovernmental organisation (NGO) *Istraživačko Dokumentacioni Centar Sarajevo* (Research and Documentation Center Sarajevo), with the assistance of other NGO’s, published the Bosnian Book of the Dead (BBD) (Tokača, 2012). The book’s four hardback volumes list almost 100,000 dead, civilian and military, by their name, their ethnicity, and when and where they died. In an early stage report before the findings were released, an international team of independent experts evaluated the project: “...the overall conclusion of our project is that the level of incompleteness and deficiency in the BBD Database is low and fairly acceptable. The database is a remarkable achievement of all those who have worked on this project” (Ball et al. (2007), p. 7).

A.2 Theoretical framework

This section provides a simple model to describe which individuals might adopt a Yugoslav identity, which guides the empirical analysis of the main text. The model is not exhaustive. Its aim is rather to build a basic intuition behind the relationship between ethnic diversity and national sentiment. I use individual theoretical analysis to generate hypotheses about the determinants of Yugoslav identification at the municipal level. My macro-level empirical framework is an aggregation of micro-level reasoning.

Assume that there are two types of national categories, Yugoslavs, Y , and non-Yugoslavs, N . A non-Yugoslav identity corresponds to the ethnic roots of an individual. The ideal non-Yugoslav has attributes a_n , while the ideal Yugoslav has attributes a_y . The identity pay-off a non-Yugoslav incurs depends on her self-image, I_n , and the extent to which her attributes, a , correspond to her category’s ideal attributes, $t_n(a_n - a)$. There might be a mismatch between who an individual is, a , and who that individual wants to be, a_n . If there is a mismatch between the two, an individual will incur a physic loss equivalent to $t_n(a_n - a)$.

Similar to Akerlof and Kranton (2002), the parameter t measures how difficult it is for individuals with different characteristics to fit in a national category. Akerlof and Kranton (2002) argue that the parameter t can be lowered through a policy of creating a sense of community. If so, a government could affect an individual’s sense of belonging to a nation through policies like education. The parameter t can likely be affected by a range of other socio-economic factors as well.

The identity pay-off of a Yugoslav is completely synonymous to that of a non-Yugoslav. An individual can choose how Yugoslav they feel, y , which is treated as a continuous variable on a unit interval. By choosing to be a Yugoslav, an individual reduces the weight on the overall identity pay-off associated with her ethnic roots. An individual can thus have multiple identities. If an individual chooses Yugoslav identity, she will incur costs, β , that are relative to the costs of remaining loyal to the ethnic roots of an individual. The costs of becoming a Yugoslav can be perceived as costs of identity switching (Caselli and Coleman, 2013), or as costs of “betraying” one’s ethnic roots (Manning and Roy, 2010). They can take several forms. They can take the form of direct monetary (opportunity) costs, like lost access to jobs allocated by ethnic criteria. They can also take the form of physic costs, like ostracism from an ethnic-based community. Finally, they can take the form of prejudice, along the lines of (Allport, 1954).

An individual thus maximises her identity pay-off by choosing how Yugoslav she feels:

$$V(y) = (1 - y)(I_n - t_n(a_n - a)) + y(I_y - t_y(a_y - a) - \beta y) \quad (5)$$

where the costs of becoming a Yugoslav are proportional to the intensity of Yugoslav feeling. If an individual does not feel as a Yugoslav at all, i.e. if $y = 0$, she will derive identity pay-off solely from her ethnic roots. Equation 5 will collapse to:

$$V_n = I_n - t_n(a_n - a) \quad (6)$$

Such individuals can be interpreted as having no belief whatsoever in the Yugoslav nation. If so, they can be characterised as ethnic nationalists. Equivalently, if $y = 1$, an individual does not feel attached to her ethnic roots at all. Instead, she firmly identifies with the Yugoslav nation. Her identity pay-off will collapse to:

$$V_y = I_y - t_y(a_y - a) - \beta \quad (7)$$

Why would ethnic diversity matter? A higher ethnic diversity could decrease the cost of identity switching, and stimulate national feeling. There are three relevant channels. First, ethnic diversity decreases the amount of resources an ethnic group can mobilise to punish its defecting member. Second, ethnic diversity decreases the amount of ethnic-specific connections a defecting individual would sever, which would otherwise involve a physic loss. Finally, according to [Allport \(1954\)](#), ethnic diversity decreases prejudice, decreasing the likelihood that an individual will be punished.

Formally, the cost of betraying one's ethnic roots depends negatively on ethnic diversity, e , and a vector of variables, z :

$$\beta = \beta(e, z) \quad (8)$$

Ethnic diversity could also have an indirect effect on the overall identity pay-off through affecting the characteristics or ethnic markers of an individual. Such ethnic markers could, for instance, take the form of the ethnic background of the spouse or partner of an individual, or the ethnic background of individual's parents. To focus on the former case, suppose that an ideal member of an ethnicity should marry or be in relationship with a member of the same ethnicity. For instance, an ideal Croat should marry another Croat, rather than say a Serb or a Slovene. In contrast, suppose that for an ideal Yugoslav the ethnicity of her spouse or partner is irrelevant.²⁵

To formalise these ideas, define $c = a_n - a$, and $d = a_y - a$. Suppose that c is an increasing function of ethnic intermarriage, i , while d is independent of ethnic intermarriage. Suppose that intermarriage in turns depends on ethnic diversity, and a vector of variables, x . Finally, assume further that both c and d depend, for simplicity, on an identical vector of other variables, f :

$$c(i, f) = c(i(e, x), f) \quad (9)$$

$$d = d(f) \quad (10)$$

Intuitively, intermarriage depends on the existence of possible matches between ethnically heterogeneous individuals, which is a supply-side issue. Intermarriage is impossible if there are no ethnically heterogeneous individuals in a society. All else given, an increase in ethnic diversity will increase the incidence of intermarriage. Even if an individual has a strong preference towards endogamy, in the extreme case of perfect ethnic heterogeneity, an individual will have to marry a member of another ethnicity unless that individual is willing to rather remain unmarried. Intermarriage will in turn increase the mismatch between the individual's attributes and the ideal attributes of a non-Yugoslav. An increase in ethnic diversity will thus increase the identity loss associated with a non-Yugoslav identity relative to a Yugoslav identity, stimulating Yugoslav sentiment.²⁶

In order to derive what determines Yugoslav sentiment, substitute equations 8, 9, and 10

²⁵Yugoslav identity, as discussed in the previous section, was not based on ethnic markers. Of course, it was initially based on the idea of South Slav unity. However, it evolved over time to be even more inclusive than that ([Djokić, 2003](#)).

²⁶These implications will hold as long as Yugoslavs cared less about the ethnicity of their spouse or partner relative to non-Yugoslavs. Of course, it is also possible to imagine that Yugoslavs valued intermarriage as an attribute. The effect would then be even more pronounced.

into equation 5, and maximise equation 5 with respect to y :

$$y = \frac{I_y - I_n + t_n c(i(e, x), f) - t_y d(f)}{2\beta(e, z)} \quad (11)$$

For the maximisation problem to have a non-negative solution on a continuous unit interval, auxiliary assumptions require that the numerator is larger than zero, and that the denominator is equal to or larger than the numerator. Ethnic diversity has an impact on Yugoslav identification through affecting the cost of becoming a Yugoslav. An increase in ethnic diversity will reduce the cost of becoming a Yugoslav, and hence stimulate Yugoslav identification. Ethnic diversity has a further impact on Yugoslav feeling through impacting the incidence of intermarriage. Subject to cost, some individuals will reject an ethnic-centred identity because they do not fit in. Instead, they will gravitate towards an identity that is not based on ethnic markers. Such individuals will identify with the wider nation rather than with a specific ethnicity. While acknowledging that other factors have an impact on national identification as well, the aim of the empirical analysis in the main text is to investigate the impact of ethnic diversity on Yugoslav self-identification.

A.3 Alternative estimators: GLM and PPML

OLS estimates might be inappropriate for two reasons. First, the relationship between ethnic fractionalisation and Yugoslav self-identification is not necessarily monotonic. Second, my estimates predict negative values of the population share of Yugoslavs. Even though the aim of the analysis is to describe, rather than to forecast, this still creates an odd problem.

There are two estimators that can be employed to assess the extent to which these two problems might bias the OLS estimates. First, [Papke and Wooldridge \(1996\)](#) propose a quasi-likelihood general linear model (GLM) for regression models with a fractional dependent variable. Their model is also flexible in dealing with alternative functional forms. Second, [Silva and Tenreiro \(2006\)](#) propose a Poisson pseudo-maximum-likelihood (PPML) estimator for regression models with a fractional dependent variable and that are characterised by non-linearities.

This section reports the GLM and PPML estimators that replicate the table 2, column 15, model. The GLM coefficient on ethnic diversity in column 1 of table 11 implies that a unit increase in ethnic fractionalisation is associated with an increase in the fraction of Yugoslavs by 0.0692 units. The PPML coefficient on ethnic diversity in column 2 of table 11 implies that a unit increase in ethnic fractionalisation is associated with an increase in the fraction of Yugoslavs by 0.0702 units. The alternative estimators imply marginal coefficients on ethnic diversity that are very similar to the coefficient on ethnic diversity provided by the OLS estimator. As such, I prefer to use OLS estimates in the main text because the interpretation of OLS coefficients is easier.

A.4 Additional control: Communist Party membership

[Burg and Berbaum \(1989\)](#) argue that Yugoslav identification signified political integration and regime support. In the main text of the paper, I have already included army presence to test their hypothesis. To test further the interpretation of Yugoslav identification as signifying regime support, I control for the population fraction of the Communist Party members in this section. However, I must restrict the analysis to Croatia due to data availability. I have not been able to find Communist Party membership records of the other regions in the archives.

Table 12 documents the OLS estimates when controlling for the Communist Party membership. Column 1 begins by showing the effect of ethnic diversity on Yugoslav sentiment in the absence of controls. In column 2, I control for the Communist Party membership. The coefficient on that variable is positive and statistically significant. Nevertheless, once I include the baseline controls in column 3, the coefficient on Communist Party membership loses significance. The coefficient

on ethnic diversity remains positive and statistically significant throughout the specifications. There is no strong evidence that communists were more likely to feel Yugoslav.

A.5 Additional control: Religious diversity

In this section, I explore the relationship between religious diversity and Yugoslav sentiment. Religious diversity is sometimes used as a proxy for ethnic diversity in the literature ([Alesina et al., 2003](#)). Of course, religious diversity might also influence nation sentiment independent of ethnic diversity.

Table 13 documents the OLS estimates when using religious diversity. I begin the analysis in column 1 by including religious diversity in the absence of controls. The coefficient on religious diversity is statistically significant, and its size is similar to the coefficient on ethnic diversity in table 2. Religiously diverse municipalities are thus associated with a greater fraction of self-declared Yugoslavs. Nevertheless, when I include both ethnic and religious diversity in column 2, the coefficient on religious diversity loses significance, while the coefficient on ethnic diversity remains significant. I find the same results when I include the baseline controls in column 3. Conditional on ethnic diversity, there is no evidence that religious diversity influenced Yugoslav identity.

A.6 Sub-samples

This section analyses the effect of ethnic diversity on Yugoslav sentiment for a variety of sub-samples. Given the inclusion of region effects in the analysis, the obvious choice is to analyse the data of each regions that constituted Yugoslavia. Table 14 documents the results. The coefficient on ethnic diversity across the regions clusters in the range of 0.041-0.077. Slovenia, however, is characterised by a lower coefficient of 0.016. The findings are statistically significant across the regions apart from Montenegro. The insignificant result for Montenegro is, perhaps, unsurprising, given that Montenegro was composed of only 20 municipalities. It might be that there is too much noise in the data, given the low amount of observations.

A.7 Settlement data

This section analyses measurement issues that might impact my results. Ethnic segregation seems to be the most important measurement problem, biasing the OLS estimates towards zero. Even if municipalities were characterised by a high degree of ethnic diversity, it might be the case that individuals within these municipalities were not exposed to the individuals of another ethnicity in practice. For example, if there was one Serbian-dominated village next to another Croatian-dominated village, the ethnic fractionalisation index will yield a high value for the municipality that contains these villages. This might be misleading, if the individuals of one ethnicity were not exposed to the individuals of another ethnicity from the other village in practice.

Measuring ethnic diversity at a more granular, settlement, level should better correspond to the actual degree of interethnic contact that individuals experience. Settlements, rather than municipalities, presumably better correspond to the daily environment of an individual, and thus to the interethnic contact that individual is exposed to.

For that matter, as [Montalvo and Reynal-Querol \(2017\)](#) argue, the relationship between ethnic heterogeneity and economic development is complex. Empirical research working with cross-country data typically finds a negative relationship. However, research that uses within-country data usually finds a positive relationship between diversity and development. This implies that ethnic diversity effects depend on the area used as the unit of observation.

A.7.1 OLS estimates with settlement data

Table 15 presents the OLS estimates of ethnic diversity when using the settlement data. Due to imperfect data, I restrict the analysis to Croatia. For comparison, I first report in column 1 the OLS estimate when I use the municipal data without any controls, which is 0.058. In column 2, I repeat the same exercise, but at settlement-level, where the coefficient on ethnic diversity increases to 0.081. Measuring the relationship at more granular level is more appropriate, yielding a higher coefficient on ethnic diversity. My baseline estimates are thus biased towards zero, providing a lower bound estimate of the true size of the ethnic diversity effect.

The drawback of using the settlement data is the lower amount of controls that I can use. The coverage of settlements by the statistical office was not as wide as it was for municipalities. Nevertheless, in column 3, I include the controls that I was able to extract. To control for economic development, I include average years of schooling, agricultural employment share and population density. Given that towns are different to rural areas across a range of characteristics, I also control for town effects. To control for potential noise in the data, I include the size of the population - many villages were extremely small, numbering only a few inhabitants. Finally, I also control for a range of relevant geographical factors, like agricultural suitability of land, terrain roughness, distance to the nearest river, distance to the coast, latitude, longitude, and the interaction between latitude and longitude. The inclusion of these control increases the coefficient on ethnic diversity to 0.12, and it remains significant at 1%.

A.7.2 IV estimates with settlement data

I now assess whether the IV strategy works when using the settlement data. The border changes, however, are still estimated for municipalities. Each settlement within the same municipality thus contains the same number of border changes. Table 16 displays the results when using the IV estimates. The second stage coefficient on ethnic diversity increases to 0.16, and is significant at the 1 per cent level. It is highly reassuring that the results are robust even when using alternative, more granular, data.

A.8 Individual-level survey data

In this section, I analyse the relationship between ethnic diversity and Yugoslav sentiment at individual-level. I was able to find data only for Croatia and Serbia that can be used for this purpose. I derive data for each of these two regions from the pre-election surveys that were conducted in 1990, in anticipation of the first democratic elections in post-war Yugoslavia.

The two surveys contain a range of questions regarding the political attitudes of the respondents, as well as their demographic and economic conditions, including their nationality. Given that the municipality of the respondents is reported as well, I am able to match this survey data with the ethnic diversity data from the population census. This allows me to test the relationship between ethnic diversity and Yugoslav sentiment at individual-level.

The results are reported in table 17. Column 1 shows the results for Croatia, while column 2 shows the results for Serbia. The relationship between ethnic diversity and Yugoslav sentiment is positive and significant in both cases. In table 18, I use the IV strategy where I instrument ethnic diversity with border changes. The first stage results are positive for both Croatia and Serbia - there is a positive association between border changes and ethnic diversity. In the second stage, I use the variation in ethnic diversity caused by border changes to predict Yugoslav self-identification. The impact of ethnic diversity on Yugoslav sentiment is positive for both Croatia and Serbia, but significant only for Serbia.

A.9 Bias from un-observables

Despite using a rich set of controls, the estimates reported in table 2 might be biased by un-observable factors correlated with selection into ethnically diverse areas. I assess the likelihood that the estimates are biased by un-observables by using the method developed by Altonji et al. (2005), and adapted to the case where the dependent variable is continuous by Bellows and Miguel (2009).

Altonji et al. (2005) argue that the selection on observables can be used to assess the potential bias from un-observables. Furthermore, the strength of selection on un-observables, relative to selection on observables, can be estimated by a ratio that would decrease the coefficient of interest to zero. The size of this ratio can in turn provide meaningful information concerning the extent of the omitted variable bias. This ratio can be estimated by using two regressions. The first regression runs a restricted model, where the estimated coefficient for the variable of interest is denoted by β^R (where R stands for restricted). The second regression runs a fully specified model, where the estimated coefficient for the variable of interest is denoted by β^F (where F stands for full). Then, the Altonji et al. (2005) ratio is calculated as $\beta^F/(\beta^R - \beta^F)$. Intuitively, the stronger is the numerator, β^F , the stronger is the effect after controlling for observables, and un-observables would have to explain more in order to reduce the coefficient of interest to zero. Concerning the denominator, $\beta^R - \beta^F$, the smaller is the difference between β^R and β^F , the estimate is affected less by selection on observables, and the stronger the selection of un-observables would need to be in order to reduce the coefficient of interest to zero.

I use two sets of controls to estimate how much stronger the effect of omitted variables would have to be, relative to observables, to attribute the entire OLS estimates to selection effects. The first specified set consists of the baseline controls in table 2, column 15, specification, while the restricted set does not include any controls. The second set adds to the specified set the additional, robustness, controls in table 5, column 7, specification in the main text, while the restricted set includes the baseline controls.

The results are presented in table 19. In the first set, the Altonji et al. (2005) ratio is 6.67, while in the second set, the ratio is 5.58. This means that selection on un-observables would have to be at least 5.58 times greater than selection on observables to attribute the entire OLS estimate to selection effects. These results suggest that the omitted variable bias would have to be very high to explain away the positive association between ethnic diversity and Yugoslav sentiment.

A.10 Alternative instruments

In this section, I experiment with alternative instruments. The relationship between border changes and ethnic diversity might be non-linear (figure 3). Column 1 of table 20 accounts for this by including a squared term of the IV. Nevertheless, the coefficient on ethnic diversity remains robust with the addition of the squared term.

In column 2, I include reconquests in the instrument. The benefit of this IV is that it varies to a greater degree across Yugoslav municipalities than the baseline instrument which excludes the reconquests (see appendix A.1 for how the baseline instrument is constructed). The cost of this alternative IV, however, is that it is more likely to be mismeasured. For example, the Great Turkish War (1683–1699) led to frequent border changes between the Habsburg and Ottoman Empires. During this war, the Habsburgs had reconquered eastern and northern areas of modern-day Croatia and Serbia, respectively. The war, however, was characterised by a multitude of offensives and counter-offensives. The front-line between the two empires were in a constant state of flux, which the historical atlases I use tend to under-report (Regan and Kaniški, 2003).

Having said this, the coefficient on ethnic diversity remains significant with the alternative IV, and its size is slightly higher than the coefficient on ethnic diversity when using the baseline

IV. I believe that the cost of the alternative instrument outweighs the benefit, and prefer to use the baseline IV in the remainder of the paper.

A.11 Spatial models

This section experiments with spatial models, as my estimates might be spatially correlated, biasing the results. Figures 1 and 2 suggest that the fraction of Yugoslavs, ethnic diversity, and border changes, tend to exhibit spatial patterns, clustered at regional-level. By including region effects, I have eliminated this source of variation between the regions. It is, however, unclear whether accounting for these spatial patterns is an appropriate research strategy. Border changes are “naturally” spatially autocorrelated, as rulers were conquering clusters of areas, which later determined the borders of Yugoslav regions. By controlling for region effects, I might have eliminated a relevant source of variation in the IV.

Nevertheless, it remains interesting to assess how the spatial nature of the data affects the results. I experiment with spatial models in table 21. In column 1, I cluster standard errors at regional level, which allows for the possibility of correlations between the error terms in the municipalities within a region. This leads to an increase in the size of the standard error on ethnic diversity from 0.010 in the baseline case (table 4) to 0.036.

It is, however, unclear whether clustering standard errors at a different level of aggregation is sufficient in dealing with spatial autocorrelation. Gibbons and Overman (2012) advocate the path taken by most recent neighbourhood effects research in spatial econometrics by including a spatial lag of the main explanatory variable. This parameter describes the influence of neighbours’ characteristics, and likely picks up the effect of the neighbours’ outcome as well. Column 2 includes the spatial lag of ethnic diversity - mean ethnic diversity of a municipality’s neighbours. The inclusion of the spatial lag leads to an increase in the coefficient on ethnic diversity to 0.223. Interestingly, the R-squared in panel B falls significantly compared to the baseline R-squared in table 4. This means that the spatial lag decreases the predictive power of the second-stage regression, and thus indicates that the spatial lag is an irrelevant control. In column 3, I include both the spatial lag and standard errors clustered at regional level, and the results are similar to those in column 2.

A.12 Reduced-form results

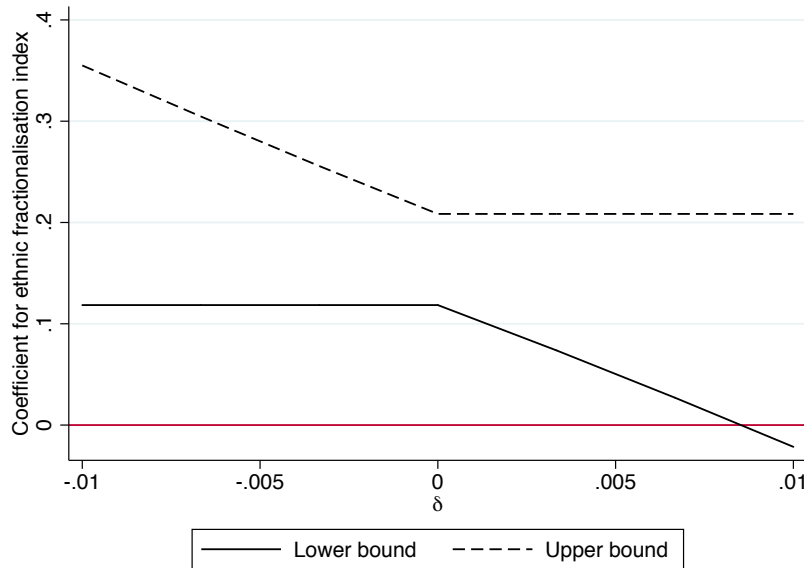
This section briefly presents the reduced-form results corresponding to the baseline 2SLS regression in column 1 of table 4. Table 22 shows the coefficient on border changes, which is highly significant and positively associated with the dependent variable - fraction of Yugoslavs. These results are consistent with the IV estimates - border changes influenced ethnic diversity, which stimulated Yugoslav sentiment.

A.13 Relaxing the exclusion restriction

The requirement of perfect exogeneity is a strict requirement. I recognize that my instrument is unlikely to satisfy the exclusion restriction perfectly, and I thus follow the approach of Conley et al. (2012) to relax it. They propose a method which allows for a direct effect of the instrument on the outcome variable - in this case, independent of ethnic diversity effects. Following their method, I assume that the potential impact of border changes on Yugoslav identification, γ , is uniformly distributed in the interval $[-\delta, \delta]$. With this method, by varying δ , I can identify the threshold at which the coefficient on ethnic diversity in the second-stage becomes statistically insignificant at the 10 per cent level.

I experiment with both a positive and negative interval of δ , as the potential direction of the impact of the instrument on the dependent variable is ambiguous. For example, a historical memory of belonging to many empires might have weakened the strength of local identities,

Figure 4: Confidence interval on ethnic diversity with relaxed exclusion restriction, [Conley et al. \(2012\)](#) method



Note: The figure shows the upper and lower bound of the 90% confidence interval of the second-stage coefficient on ethnic diversity, using table 4, column 1, model. It follows the union of confidence intervals approach in [Conley et al. \(2012\)](#), which allows for a direct effect of border changes on Yugoslav identification, assuming this is uniformly distributed in the interval $[-\delta, \delta]$.

stimulating Yugoslav sentiment. On the other hand, it is also possible that historical memory of belonging to many empires antagonized the local population, strengthened their local identities, and thus negatively impacted Yugoslav sentiment.

Figure 4 relaxes the exclusion restriction, and identifies the threshold of $\delta = 0.0086$ at which the coefficient on ethnic diversity in the second-stage becomes statistically insignificant at the 10 percent level. Put alternatively, as long as the direct effect of border changes on Yugoslav identification is smaller than 0.0086, the coefficient on ethnic diversity remains significant at the 10 per cent level. For that matter, if $\delta < 0$, the confidence interval on the coefficient on ethnic diversity moves further from zero, relative to the baseline IV estimate in table 4. In other words, if border changes had a negative impact on Yugoslav identification, then the IV coefficient underestimates the true effect of ethnic diversity on Yugoslav sentiment.

To gauge the magnitude of the threshold at which the coefficient on ethnic diversity in the second-stage becomes statistically insignificant, I estimate the reduced-form effect of border changes on Yugoslav sentiment in appendix A.12, which is 0.0099. The impact of the instrument on Yugoslav identification ($\delta = 0.0086$) would thus have to be equivalent to 87 per cent of the overall reduced-form effect to render the validity of the 2SLS results insignificant. The analysis thus reveals that the instrument would have to deviate very far from the exclusion restriction to make my results insignificant.

A.14 Mediation analysis

Estimating the intermarriage mechanism is difficult. This requires identifying the causal effect of the treatment (ethnic diversity) on the final outcome (Yugoslav sentiment), as channelled the mediator (intermarriage). Moreover, this also requires estimating the direct effect of ethnic diversity on Yugoslav identity, using a single instrument (border changes). [Dippel et al. \(2017\)](#) propose a framework to deal with this problem. Their framework estimates the effect of the endogenous treatment on the final outcome (total effect), while decomposing this total effect

into a “direct effect” and an “indirect effect”, which runs through the mediator. The advantage of this method over the standard IV is that, with a single instrument, the causal effect of a mediator on the final outcome can also be estimated.

To provide more details: the first step of the procedure is to estimate the effect of the treatment (ethnic diversity) on the final outcome (Yugoslav sentiment), while instrumenting the treatment with an IV (border changes). This step yields the “total effect” of the ethnic diversity on Yugoslav identity. The second step is to estimate the effect of the ethnic diversity on the mediator, again using border changes as an instrument. The final step involves another IV estimation where the effect of the mediator on Yugoslav sentiment is estimated, conditional on the treatment. In short, the framework provides a mediation effect which is essentially a linear combination of coefficients from three different regression. The full details can be found in [Dippel et al. \(2017\)](#).

The mediation analysis is presented in table 23. The total effect estimates that a unit increase in the ethnic fractionalization index increases the fraction of Yugoslavs by 0.104 units. Under the typical IV assumptions, this effect is causal. The direct effect estimates that only 0.0298 units of this increase are due to ethnic diversity itself. However, the indirect effect estimates that 0.059 units of this increase are caused by intermarriage as a mediating factor. This implies that intermarriage explains 57% of the total effect of ethnic diversity on Yugoslav sentiment.

A.15 The meaning of Yugoslav identity

In this section, I briefly outline the meaning and content of Yugoslav identity. I provide anecdotal evidence by focusing on the work and thoughts of the leading proponents of a supranational Yugoslav culture - the unifying glue of a Yugoslav nation. It is difficult to exactly define what Yugoslav identity meant. Yugo-Slavism was never clearly defined. It was interpreted in a variety of ways ([Djokić, 2003](#)). However, this does not mean that Yugoslav identity had no meaning or content. On the contrary.

[Wachtel \(1998\)](#) argues that attempts of Yugoslav nation building can be divided into three basic categories concerning how a national culture or identity should be created, and on what content should it be based upon. First, according to the earliest Yugoslav view, the South Slavs should have adopted a modified form of Serbian culture. The Serbs were the largest ethnic group, and were the first in modern times to create an independent state. The proponents of the second view thought that Yugoslav identity should be synthetic. The ideal Yugoslav culture should adopt the best elements from the separate South Slavic cultures to create an overreaching synthetic Yugoslav culture. Both views, however, did not imply that separate South Slavic cultures should cease to exist. But, rather, that they would be made less relevant with the emergence of a new Yugoslav culture. Separate South Slavic cultures would take a role of local customs and traditions. The third view that emerged in the post-war period was that common socialist features should be the unifying force of the country (section 2).

[Wachtel \(1998\)](#) argues that, of the above-mentioned views towards the construction of the Yugoslav identity, the second view of a synthetic culture was the prominent one during the 20th century. To get a better sense of what synthetic culture meant, it is useful to concentrate on the work and thoughts of the leading proponents of it. [Wachtel \(1998\)](#) argues that the two most publicly visible and famous proponents of a supranational synthetic Yugoslav culture were Ivan Meštrović, a sculptor, and Ivo Andrić, a Nobel prize-winning author, both of whom were of Croatian origin.

During the interwar period, Meštrović became a world-famous artist and the leading spokesman for a new Yugoslav culture. [Banac \(1984\)](#) reports that those who knew Meštrović’s views referred to him as “the prophet of Yugo-Slavism”. Meštrović believed that sculpture offered an ideal form for expressing a new Yugoslav synthesis. Sculpture is more accessible to general audience than high literature. Moreover, by sculpting figures of mostly Serb epic poetry, South Slavic identities could be melded. Sculpture was largely non-existent in the Serbian

Orthodox-based culture, which prohibits three dimensional figures of divinity. On the other hand, sculpture was well-developed in the renaissance Dalmatian towns.

Meštrović entered the Yugoslav public scene and became the best-known domestic artist with his controversial exhibition at the Rome Exposition of 1911. There, he created a sculpture of Serbian folk-legend hero Prince Marko. Meštrović described the statue of Prince Marko as symbolising the Yugoslav people, “with its gigantic and noble heart” (Banac, 1984). At the Rome exhibition, Meštrović also created a model of the temple of Kosovo, which holds a central location in the Serbian national myth as a birthplace of the medieval Serbian Kingdom. His temple of Kosovo combined Catholic and Orthodox architectural elements. The visual symbolism was obvious. The aim was to create an architectural synthesis of Yugoslav identity. Temple of Kosovo would be something to Yugoslavs what, say, Louvre is to the French (Wachtel, 1998). In addition to his visual representations of the Yugoslav nation, Meštrović wrote poetry speaking of a “Yugoslav race” (Banac, 1984).

As additional anecdotal evidence, consider the work of Ivo Andrić. His work symbolised, perhaps, the postwar Yugoslav literary canon the most. During WWII he wrote his two most famous novels, *Na Drini Čuprija* (The Bridge on the Drina) and *Travnička Kronika* (Bosnian Chronicle). Upon their release in the aftermath of WWII, the works were almost instantly proclaimed classics of the Yugoslav literature (Wachtel, 1998). His two great novels are set in multicultural and multi-ethnic Bosnia. They chronicle the difficult historical interaction of the ethnicities of Yugoslavia, while holding hope for a supranational union that might bind them together. For example, the bridge in the *The Bridge on the Drina* is often interpreted, rightly or wrongly, as symbolising Yugoslavia (Wachtel, 1998). Like the bridge, Yugoslavia was to be a mediator between Eastern and Western cultural traditions, connecting and blending both in a unique synthesis.

There were, of course, other artists that promoted Yugoslav identity. A prominent institution whose aim was to promote and develop Yugoslav identity was the literary magazine *Književni Jug* (Literary South). Its explicit aim was to lay ground for a future Yugoslav literature. Its duration was short lived (1918-1919), but many of the authors that are considered today as some of the greatest authors of the successor states of Yugoslavia published there. These include the Slovenian Ivan Cankar, the Croatian Vladimir Nazor, Tin Ujević, and Ivo Vojnović, the Serbian Sima Pandurović, and the Bosnian Aleksa Šantic (Milisavac, 1971).

In terms of popular culture, Yugoslav rock music is considered to be a crucial foundation of supranational identity in late Yugoslavia, especially in the youth. Ramet (1992) argues that rock music in Yugoslavia was seen as by many of its purveyors as pan-Yugoslav, a force that brought people together and created ties of mutual acceptance. It created a shared cultural space and common reference points, superseding ethnic boundaries (Perković, 2011). Ramet (1992) argues that, symptomatically, many of the leading figures of the Yugoslav rock scene emphasised that they were “Yugoslavs”, as opposed to say Serbs or Slovenes.

It was, however, not only rock music that served as a purveyor and constructor of Yugoslav feeling. Pop-folk, which reached extremely high levels of popularity during the 1980s (and remains equally popular today in Yugoslav successor states, if not more), also mattered. Consider the example of Fahreta Živojinović, better known as Lepa Brena. Since her pop-folk career began in the early 1980s, Lepa Brena became arguably the most popular singer of Yugoslavia. She still maintains a strong popularity in Yugoslav successor states. In 1989, Lepa Brena released the controversial song *Jugoslavenka* (Yugoslavian (girl)), which is unashamedly pro-Yugoslav. It remains one of her most popular songs. While there were plenty of popular song earlier celebrating Yugoslavia and its identity, many were implicitly or explicitly glorifying the socialist regime. As such, Ramet (1992) describes some of these “patriotic” songs as sycophantic. Yugoslavian by Lepa Brena, on the other hand, seems genuine. It was written at the time of rising ethnic nationalism. Pro-Yugoslav songs during the late 1980s were strongly damaging to a singer’s career (Perković, 2011). In the song, Lepa Brena relates features of a beautiful Yugoslav girl

with the geographic and other features of Yugoslavia:

Where are you from, pretty girl
who gave birth to the blue eye
who gave you the golden hair
who made you so passionate

Chorus x2

My eyes are Adriatic sea
my hair is Panonian wheat
wistful is my Sloven soul
I'm Yugoslavian

Where are you from, pretty girl
where did you grow, spring flower
where free sun is warming you
when you dance so seductively

Chorus 2x

Where are you from, pretty stranger (girl)
where have you been stealing sun's shine
where were you drinking honey wine
when your kiss is so sweet

Chorus 2x

It is, perhaps, not surprising that Lepa Brena still explicitly identifies herself as “Yugoslav”, and feeling “Yugonostalgic”. In her own words in 2008: “If someone has a right to declare themselves as Croat or Serb, I too have a right to declare myself as Yugoslav. Yugoslavia was specific in many ways... Similar to [former] Yugoslavia lives united Europe [today], which is a civilised society. I admit, I am Yugonostalgic, and I think you shouldn't close doors to other people. You have to respect the right of people with another colour of skin, religion, and diversity...” ([Index.hr](#), 2008, December 8) (author's translation, squared brackets inserted by author).

Table 11: GLM and PPML estimates, dependent variable: fraction of Yugoslavs

	(1) GLM	(2) PPML
Ethnic fractionalisation	2.7127*** (0.187)	2.5970*** (0.177)
Output p.c.	0.0007 (0.016)	0.0007 (0.015)
Population density	0.0000 (0.000)	-0.0000 (0.000)
Avg. years of schooling	0.4107*** (0.069)	0.3870*** (0.066)
Social sector labour	0.1811 (0.339)	0.1825 (0.324)
Youth labour actions	-0.6430 (0.396)	-0.6027 (0.380)
WWII partisan veterans	11.1942 (11.916)	10.7425 (11.349)
Monuments to anti-fascism	-0.1301 (0.091)	-0.1254 (0.087)
Fascist terror	2.8319*** (0.632)	2.7390*** (0.604)
Public goods exp. p.c.	-0.0060 (0.011)	-0.0052 (0.011)
Federal aid	-0.0557 (0.108)	-0.0467 (0.104)
1980's generation	-2.5976 (3.423)	-2.3216 (3.300)
Army presence	0.2616*** (0.098)	0.2495*** (0.093)
Region effects	Yes	Yes
Observations	434	434
R-squared	n.a.	0.639

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 12: OLS estimates with Communist Party membership, dependent variable: fraction of Yugoslavs

	(1)	(2)	(3)
Ethnic fractionalisation	0.0579*** (0.007)	0.0537*** (0.008)	0.0585*** (0.009)
Communist Party members		0.1630*** (0.058)	0.0654 (0.069)
Baseline Controls	No	No	Yes
Observations	102	102	102
R-squared	0.503	0.529	0.640

Notes: The analysis is restricted to Croatia due to data availability. Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 13: OLS estimates with religious diversity, dependent variable: fraction of Yugoslavs

	(1)	(2)	(3)
Religious fractionalisation	0.0787*** (0.008)	-0.0190 (0.030)	0.0005 (0.029)
Ethnic fractionalisation		0.1087*** (0.031)	0.0790*** (0.029)
Baseline controls	No	No	Yes
Observations	434	434	434
R-squared	0.242	0.321	0.555

Notes: Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 14: OLS estimates, sub-samples, dependent variable: fraction of Yugoslavs

	(1) Croatia	(2) Bosnia	(3) Montenegro	(4) Serbia	(5) Slovenia
Ethnic fractionalisation	0.0595*** (0.009)	0.0424*** (0.011)	0.0414 (0.075)	0.0771*** (0.020)	0.0162** (0.008)
Baseline controls	Yes	Yes	Yes	Yes	Yes
Observations	102	109	20	143	60
R-squared	0.638	0.728	0.834	0.576	0.775

Notes: Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 15: OLS estimates with settlement data, dependent variable: fraction of Yugoslavs

	Municipality level (1)	Settlement level (2) (3)	
Ethnic fractionalisation	0.0579*** (0.007)	0.0814*** (0.014)	0.1201*** (0.023)
Controls	No	No	Yes
Observations	102	6,295	6,290
R-squared	0.503	0.027	0.048

Notes: The analysis is restricted to Croatia due to data availability. Controls are average years of schooling, agricultural employment share, town effects, population density, population, agricultural (wheat) suitability, terrain roughness, distance to the coast, distance to the nearest river, longitude, latitude, and the interaction between longitude and latitude. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 16: 2SLS estimates with settlement data, dependent variable: fraction of Yugoslavs

(1)	
Panel A: first stage outcome - ethnic fractionalisation	
Border changes, 1421-1816	0.0444*** (0.003)
Controls	Yes
R-squared	0.3355
Panel B: second stage outcome - fraction of Yugoslavs	
Ethnic fractionalisation	0.1706*** (0.024)
Controls	Yes
Observations	6,290
Centered R-squared	0.0406
Sanderson-Windmeijer first stage F-test	202.75***

Notes: The analysis is restricted to Croatia due to data availability. Controls are average years of schooling, agricultural employment share, town effects, population density, population, agricultural (wheat) suitability, terrain roughness, distance to the coast, distance to the nearest river, longitude, latitude, and the interaction between longitude and latitude. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 17: Individual-level data, dependent variable: Yugoslav self-identification

	(1) Croatia	(2) Serbia
Ethnic fractionalisation	2.0587*** (0.403)	36.6378*** (8.808)
Settlement type	Yes	Yes
Gender	Yes	Yes
Age	Yes	Yes
Education	Yes	Yes
Occupation	Yes	Yes
Communist party memebrship	Yes	Yes
Income	No	Yes
Observations	2,440	1,314
Number of municipalities	12	13
Pseudo R-squared	0.1014	0.1514

Notes: The unit of observation is the individual. The coefficients are derived from a logit estimation. Settlement type includes three categories in Croatia, and four categories in Serbia. Education includes five categories in Croatia, and six in Serbia. Occupation includes ten categories in Croatia, and nine in Serbia. Communist Party membership includes three categories in Croatia, and two in Serbia. Income includes six categories in Serbia. Ethnic fractionalisation is measured at municipal-level. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Data: Survey data for Croatia is taken from [Siber \(1997\)](#), and survey data for Serbia is from [Goati \(1998\)](#).

Table 18: IV estimates with individual-level data, dependent variable: Yugoslav self-identification

	(1) Croatia	(2) Serbia
Panel A: first stage outcome - ethnic fractionalisation		
Border changes, 1421-1816	0.0379*** (0.004)	0.0352*** (0.003)
Settlement type	Yes	Yes
Gender	Yes	Yes
Age	Yes	Yes
Education	Yes	Yes
Occupation	Yes	Yes
Communist party memebrship	Yes	Yes
Income	No	Yes
Panel B: second stage outcome -self-identified Yugoslavs		
Ethnic fractionalisation	0.0017 (1.175)	38.4917*** (14.242)
Settlement type	Yes	Yes
Gender	Yes	Yes
Age	Yes	Yes
Education	Yes	Yes
Occupation	Yes	Yes
Communist party memebrship	Yes	Yes
Income	No	Yes
Observations	2,440	1,314
Number of municipalities	12	13

Notes: The unit of observation is the individual. Settlement type includes three categories in Croatia, and four categories in Serbia. Education includes five categories in Croatia, and six in Serbia. Occupation includes ten categories in Croatia, and nine in Serbia. Communist Party membership includes three categories in Croatia, and two in Serbia. Income includes six categories in Serbia. Ethnic fractionalisation is measured at municipal-level. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data: Survey data for Croatia is taken from [Siber \(1997\)](#), and survey data for Serbia is from [Goati \(1998\)](#).

Table 19: Bias from un-observables, [Altonji et al. \(2005\)](#) method, dependent variable: fraction of Yugoslavs

Controls in the restricted set	Controls in the full set	(1)
None	Baseline controls	6.67
Baseline controls	Additional controls	5.58

Notes: Following the methodology of [Altonji et al. \(2005\)](#), the table reports the strength of selection on unobservables, relative to observables, that is required to attribute the entire OLS estimates of the relationship between ethnic diversity and Yugoslav identification to selection effects. Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. Additional controls are fraction of minorities, fraction of Muslims, duration of Habsburg rule, agricultural (wheat) suitability, coast, and geographic controls. These controls are identical to those in column 7 of table 5.

Table 20: 2SLS estimates with alternative instruments

	(1)	(2)
Panel A: first stage outcome - ethnic fractionalisation		
Border changes, 1421-1816	0.0759*	
	(0.039)	
Border changes, 1421-1816 - squared	-0.0005	
	(0.007)	
Border changes, 1421-1816 (including reconquests)		0.0256***
		(0.005)
Baseline controls	Yes	Yes
R-squared	0.403	0.370
Panel B: second stage outcome - fraction of Yugoslavs		
Ethnic fractionalisation	0.1343***	0.1608***
	(0.027)	(0.028)
Baseline controls	Yes	Yes
Observations	434	434
Centered R-squared	0.479	0.387
Sanderson-Windmeijer first stage F-test	29.01***	26.24***

Notes: Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 21: 2SLS estimates with spatial models

	(1)	(2)	(3)
Panel A: first stage outcome - ethnic fractionalisation			
Border changes, 1421-1816	0.0729*** (0.007)	0.0280*** (0.008)	0.0280* (0.013)
Spatial lag - Ethnic fractionalisation		0.7651*** (0.053)	0.7651*** (0.049)
Region-clustered standard errors	Yes	No	Yes
Baseline controls	Yes	Yes	Yes
R-squared	0.403	0.625	0.625
Panel B: second stage outcome - fraction of Yugoslavs			
Ethnic fractionalisation	0.1353*** (0.036)	0.2230*** (0.077)	0.2230*** (0.085)
Spatial lag - Ethnic fractionalisation		-0.1091* (0.063)	-0.1091* (0.064)
Region-clustered standard errors	Yes	No	Yes
Baseline controls	Yes	Yes	Yes
Observations	434	434	434
Centered R-squared	0.476	0.224	0.224
Sanderson-Windmeijer first stage F-test	106.48***	10.90***	4.62*

Notes: Baseline controls are output p.c., population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 22: Reduced-form results, dependent variable: fraction of Yugoslavs, OLS

	(1)
Border changes, 1421-1816	0.0099*** (0.002)
Observations	434
R-squared	0.442
Controls	Yes

Notes: Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence, and region effects. These controls are identical to those in column 15 of table 2. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 23: Mediation analysis, dependent variable: fraction of Yugoslavs

	(1)
Total effect of ethnic fractionalization	0.1037*** (0.026)
Direct effect of ethnic fractionalization	0.0298* (0.017)
Indirect effect of ethnic fractionalization	0.059** (0.029)
Baseline controls	Yes
Observations	109
F-statistic for excluded instruments	
- First stage, step one (treatment regressed against instrument)	49.473
- First stage, step two (mediator regressed against instrument, conditional on treatment)	11.07

Notes: The mediation analysis is conducted following the framework of [Dippel et al. \(2017\)](#). The indirect effect of ethnic diversity operates through the channel of intermarriage (mediator). The instrument is border changes during 1421-1816. The analysis is restricted to Bosnia-Herzegovina since the intermarriage data is not available for other regions. Baseline controls are output p.c, population density, average years of schooling, fraction of social sector employment, population fraction of youth labour actions, population fraction of WWII partisan veterans, sites containing monuments to anti-fascism, population fraction of individuals exposed to WWII fascist terror, expenditure on public goods p.c., federal aid, population percentage of the early 1980's generation, sites containing army presence. These controls are identical to those in column 15 of table 2. Standard errors in parentheses. Note however, that they do not hold any conventional meaning, and should be interpreted with caution. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.