Intergroup Contact and Nation Building: Evidence from Military Service in Spain*

Julio Cáceres-Delpiano† Antoni-Italo De Moragas‡
Gabriel Facchini§ Ignacio González¶

Forthcoming at Journal of Public Economics

Abstract

We study the long-term effects of intergroup contact on nation building by exploiting a national lottery that randomly allocated conscripts to different military areas across Spain. For men born in regions that feature a weak Spanish identity, we find that being assigned to military service in a different region substantially increases self-identification as Spanish, increases the likelihood of voting in national elections and reduces the probability of voting for a regionalist party. Moreover, in support of intergroup contact as the main mechanism behind these results, we find that movers are more likely than nonmovers to have friends from other regions.

JEL Classification: D02, D71, J15
Keywords: Nation Building, Intergroup Contact, Military Service.

*The authors thank the editor, Maria Petrova, and three anonymous referees for their very valuable comments and suggestions. They also thank Laia Balcells; Samuel Bazzi; Federico Curci; Monica Deza; Juanjo Dolado; Ruben Durante; Lucie Gadenne; Libertad González; Gianmarco León; Mónica Martínez-Bravo; Santiago Sánchez Pagés; Charles-Louis Sidois; Marga Torre; and the seminar participants at Universidad Carlos III de Madrid, EALE Conference 2020, and the Econometric Society Virtual World Congress 2020. Cáceres-Delpiano gratefully acknowledges financial support from the Spanish Ministry of Education (Grant ECO2019-00419-001), the Spanish Ministry of Economy and Competitiveness (MDM 2014-0431) and the Comunidad de Madrid, SPAIN (Grant CAM-EPUC3M11). De Moragas acknowledges the support of the Spanish Ministry of Education (grant PGC2018-099415-B-100 MICINN/FEDER/UE) and Fundación Ramón Areces. Facchini gratefully acknowledges financial support from the General Secretariat for Research-Government of Catalonia (SGR2017-1301) and the Spanish Ministry of Education (PID2019-104619RB-C43).

†Department of Economics, Universidad Carlos III de Madrid. jcaceres@eco.uc3m.es
‡Department of Economics, CUNEF. antoni.demoragas@cunef.edu
§Department of Applied Economics, Universitat Autònoma de Barcelona. gabriel.facchini@uab.cat
¶Department of Economics, American University. ignaciog@american.edu

1 Introduction

In a world where the number of ethnic groups outweighs the number of countries, multiethnic societies are an inescapable reality and are likely to increase through migration (Putnam, 2007). Recent research suggests that ethnic diversity might be beneficial for economic development by enhancing productivity and innovation, particularly in local economies.[1] However, ethnic diversity also poses several challenges for societies as it can generate institutional instability, an inefficient provision of public goods and higher likelihood of conflict.[2] When ethnically diverse societies succeed in developing policies that build a common national identity, such negative effects can be ameliorated (Miguel, 2004). Therefore, understanding how countries can create this collective identity, a process known as nation building, becomes a crucial question, especially in the context of increasing diversity.

In this paper, we study whether intergroup contact can promote nation building. Traditionally, two hypotheses about intergroup interaction have been postulated. First, the contact hypothesis states that contact should reduce the problems of stereotyping, prejudice, and discrimination, thereby fostering interethnic tolerance and solidarity and favoring a shared national identity (Allport, 1954). In contrast, the conflict hypothesis argues that higher intergroup contact can increase the competition for limited resources and thus reinforce mistrust between different groups, thereby making the emergence of a shared national identity more difficult (Blumer, 1958).

From an empirical point of view, identifying the effect of contact is a common concern because contact can be endogenous. To address this issue, we exploit the quasi-experimental variation caused by a national lottery that allocated conscripts to different military areas in Spain, a country with historically strong regional cleavages. From 1987 to 1991, conscripts in Spain entered a national lottery that determined where they would complete their military service. An individual's destination camp depended on his province of residence and date of birth. Consequently, people from the same province in Spain fulfilled their military service duties in different parts of the country, thereby increasing their exposure to colleagues and civilians from other regions. Our empirical strategy relies on comparing the national identity and political behavior of individuals with different likelihoods of completing their military service in a region other than their own.

We use three waves (2013, 2015 and 2017) of the General Spanish Social Survey conducted by the Centre for Sociological Research (abbreviated CIS in Spanish). This survey is well suited for our analysis since the survey contains self-reported preferences on national identity and political behavior and variables that define the risk of being assigned to a particular region. We combine the surveys with the military service lottery results for the years 1987 to 1991.

Strikingly, twenty-five years later, comparable men from regions with a weak Spanish

[1] See Ottaviano and Peri (2005), Ottaviano and Peri (2006), Lee (2015), Alesina, Harnoss and Rapoport (2016), Montalvo and Reynal-Querol (2021), among others.

[2] See Easterly and Levine (1997), Alesina, Baqir and Easterly (1999), La Porta et al. (1999), Montalvo and Reynal-Querol (2005), Michalopoulos and Papaioannou (2016), Desmet, Gomes and Ortuño-Ortín (2020), Arbatlı et al. (2020), among others.

identity who were assigned to military service outside their region were more likely than those who stayed within their region to identify themselves as Spanish. More precisely, we find that being assigned to a camp outside their region increases their self-reported identity as Spanish by approximately seventeen percentage points. However, the effect is negative and not statistically different from zero among men born in regions with a strong Spanish identity. We also find effects on political behavior. First, we find that individuals assigned to a camp outside their region are approximately eighteen percentage points more likely to participate in national elections. We read this finding as a sign of greater engagement with Spanish society and institutions. Second, these individuals are also fourteen percentage points less likely than those assigned to their own region to vote for a regionalist party. This effect on political behavior is consistent with our findings on identity formation.

To rule out the usual concern that our source of variation might be confounded with other unobserved factors, we study women born from 1987-1991 and for whom military service was not mandatory. When we replicate the same exercise for women we find a small and non-statistically significant effect.

Finally, we use evidence on friendship formation to further explore the intergroup contact hypothesis. We find that individuals assigned to another region are sixteen percent more likely to report having someone from another region in their inner social circle. This finding is consistent with peers affecting an individuals' national sentiment through intergroup contact.

The case of Spain is uniquely well suited to studying the effect of intergroup contact on nation building. Spain is a multilinguistic country with historically significant ethnolinguistic cleavages. Approximately twenty percent of the population speaks a language different from Spanish at home (Pew Research Center, 2019). Moreover, the existence of historical, regional and peripheral nationalist movements—a further manifestation of such cleavages— continues to shape the political landscape in Spain.

Our paper joins a large and growing empirical literature that addresses the effects of intergroup contact. A group of papers has explored how individuals who have contact with minorities are more prone to develop either future interracial interactions or tolerant attitudes and affirmative action principles (Boisjoly et al., 2006; Scacco and Warren, 2018; Merlino, Steinhardt and Wren-Lewis, 2019; Carrell, Hoekstra and West, 2019; Dahl, Kotsadam and Rooth, 2020). More specifically related to our question, Okunogbe (2018), Bazzi et al. (2019) and Kukic (2019) study how interethnic exposure contributes to nation building for different groups and institutional settings. Different from these papers, our work shows the long-term effects of a short period of exposure. Since military service in Spain was mandatory for the entire male population, our study also has advantages in terms of external validity.

Our empirical analysis also relates to a recent body of work that studies how specific policies can shape national identity or contribute to community building (Aspachs-Bracons et al., 2008; Bandiera et al., 2018; Alesina, Reich and Riboni, 2020). A salient aspect of our work compared to this literature is that nation building is not an outcome of a policy specifically devised for this purpose but an unintended consequence of its design. Thus,

our work is related to several papers that show how certain events have the concomitant effect of fostering nation building (Dell and Querubin, 2017; Depetris-Chauvin, Durante and Campante, 2020; Dehdari and Gehring, 2019; Casas, Curci and De Moragas, 2020).

This paper also relates to the literature that studies the effect of military service on several outcomes (Angrist, 1990; Galiani, Rossi and Schargrodsky, 2011; Card and Cardoso, 2012; Cáceres-Delpiano, 2019), specifically, to papers that study the formation of political attitudes and beliefs (Erikson and Stoker, 2011; Samii, 2013; Navajas et al., 2020; Fize and Louis-Sidois, 2020). However, unlike this literature, we do not examine the effect of the military service itself but a particular feature of military service in Spain: the regional allocation of conscripts. This approach helps us study the long-term effects of interpersonal contact and to uncover potential heterogeneity in the effects of military service.

Finally, our paper is closely related to the contemporaneous work of Bagues and Roth (2020) that also exploits the allocation of conscripts in the Spanish military service to study the effect of intergroup exposure on a set of outcomes, including identity. Their results on national identity are consistent with ours. However, unlike us, they do not find significant effects on the likelihood of voting for regionalist parties. A possible explanation for this discrepancy is the timing of the fieldwork of our respective surveys. Between the fieldwork of our surveys (April 2013 - June 2018) and theirs (December 2019 - February 2020), several convulsive political events occurred that might have shifted individuals' political behavior (i.e., the jail sentences of several members of the Catalan government and secessionist political leaders in October 2019).

2 Institutional background

2.1 Regional and national identities in Spain

A central issue in Spanish politics is the so-called "national question", that is, how to accommodate regional identities in a common national project (Muñoz, 2009). The tension between national unity and regional diversity that has characterized the political landscape in Spain is well reflected in the ambiguity of the Spanish Constitution, which recognizes the right to the autonomy of "nationalities" –which supposedly have a historical regional identity– as well as "regions" –which supposedly lack such identity– but refrains from defining the scope of this right for each type of territory. Despite lacking a list of territories that are considered to be historical nationalities, as opposed to simple "regions", three territories–Catalonia, the Basque Country and Galicia– have been considered the historical nationalities in Spain (Moreno, 2002). These regions also have a language other than Spanish (Catalan, Basque and Galician), secessionist and peripheral nationalist parties with Parliamentarian representation and a large share of the population with strong feelings of regional identity.

In addition to these regions, Navarre also features a large share of the population who do not self-identify as Spanish. This fact is probably explained by the large fraction of the Navarre population who shares a Basque identity, a feeling that is particularly strong in

the areas that are close to the border with the Basque Country. Actually, according to the nationalism saliency index, Navarre, together with the Basque Country and Catalonia, is the region where peripheral nationalism is most salient (Amat, 2012).

In the remainder of this article, we refer to the aforementioned historical nationalities plus Navarre as regions with weak Spanish identity (WSI), while the rest of the regions are considered to have a strong Spanish identity (SSI).[3]

2.2 The military national lottery

Spain abolished mandatory military service in 2001, thereby putting an end to 230 years of compulsory military service. Although there was always some degree of randomization in the recruitment process, between 1987 and 1991 the recruitment was implemented through a public national lottery, which we use as a quasi-experimental device.

During this period, the conscription process consisted of three consecutive phases. First, a wide-reaching registration occurred at the municipal level every year; in this process, individuals had to register themselves approximately two years before their conscription, which resulted in provisional lists of recruits. Individuals who deferred military duties for academic studies and other reasons were excluded from these lists.[4] Second, the potential recruits had to receive a medical checkup, and a final list naming those who were sufficiently fit to serve was publicly posted. The third phase was the allocation of conscripts by using the national lottery.

The lottery occurred approximately one year before conscription. This lottery consisted of selecting a single ball, out of 366, with a specific date. Starting with those whose birthday was immediately after the selected date, an algorithm allocated the conscripts to different destinations by sequentially filling the military areas of the army, followed by those of the navy, and ending with the military areas of the air force. There was a total of nineteen military destinations at the time: ten military areas for the army; four areas for the air force; and five areas for the navy, including the navy's central jurisdiction (see figure A.7 in the online Appendix for maps containing the military areas). The final allocation outcome consisted of different birth date ranges by province (sub-regional administrative unit) that determined the destination of each conscript. Figure A.8 in the online Appendix shows the example of the lottery outcome for the province of Barcelona.[5]

The allocation algorithm followed personnel goals that were predefined at the national

[3]Note that the Spanish Constitution uses the term "Autonomous Communities" to institutionally configure all the regions, regardless of their historical character. For the sake of simplicity, and given that we use terms as "regional identity" or "regionalist parties", throughout the text we simply use region as a synonym of Autonomous Community.

[4] Five circumstances could be adduced to defer military service: i) academic studies; ii) family economic support; iii) having another sibling in the army; iv) residing abroad; or v) being elected to a public position by a public vote. Most of the deferrals were related to pursuing academic studies. In 1990, for example, from a cohort of approximately 360,000 individuals, 160,000 were temporarily excluded, among which 77% were excluded for academic reasons. Apart from these circumstances, a fraction of 0.4% invoked conscientious objection in 1990.

[5]The algorithm enabled a fast allocation of conscripts. For example, according to news reported in 1987, the algorithm took just three minutes to assign all the individuals to their respective military areas (LaVanguardia, 1987).

level. Among these goals, every year, the governmental authority determined the proportion of individuals who would serve in the military area to which their place of residence belonged. During the years of the national lottery, the majority of conscripts would serve in their military area of residence (e.g., in 1987 the government set a target of 60% of individuals that should serve in their area (BOE, 1987)). If this target was reached during the allocation process, then the remaining conscripts were assigned to other military areas. Although this target varied across regions depending on their population size, it was the same for all the provinces belonging to a region.[6] In our econometric model, we exploit the variation induced by the lottery and the regional targets. Conditional on annual national military cohort goals (captured by year of birth fixed effects) and regional targets (captured by region fixed effects), being sent outside one's own region of residence was random.[7]

3 Data and empirical strategy

3.1 Data

Sample construction: We use three waves of the General Spanish Social Survey (ESGE) conducted by the Center for Sociological Research (CIS) in Spain. The waves were collected in 2013 (study 2975), 2015 (study 3123) and 2017 (study 3201) and are representative of the Spanish population older than eighteen years. For each individual, we have information on their sociodemographic characteristics (month, year and province of birth and education), political attitudes (national identity and voting), partner characteristics and also – but only for the wave of 2013– friend characteristics.

The combined sample includes 14,163 individuals born in Spain. We select the subsample of 1,428 individuals who were exposed to the public national lottery, which are the individuals who were born between September 1968 and December 1973. Additionally, we drop 97 observations for which information on identity or voting is missing. Moreover, because our national identity question below opposes Spanish identity to the identity from the region of residence, we exclude from the analysis the individuals who reside at the time of the survey in a region different from the region of their birth (183 observations) for whom the question might fail to capture the proper regional identity. In Table 3 of Section 3, we show that the results are virtually the same when we also include these individuals. Furthermore, although military service could affect the place of residence at the time of the survey, in Table A.3 of the online Appendix we show that this was not the case.

We merge the resulting database with the allocation induced by the lottery. The lottery was based on the province of residence at the age of registration and the date of birth. We use the province of birth as a proxy for the province of residence at the age of registration. We validate this proxy with the 1991 Sociodemographic Survey conducted by the Spanish

[6] In general, small regions such as Baleares or Canarias had a higher share of conscripts who stayed in their own region due to a lower supply of local conscripts to fill the local military positions.

[7] Note that we use the year of birth instead of the year of the lottery. This is for two reasons. One, we lack information on the year that the individuals participated in the lottery. Second, since individuals can postpone the lottery, the actual year of participating can be endogenous.

Statistical Office, and we find that 89% of the individuals born during the years that we consider resided in the province of birth at the age of 17 (i.e., the age that corresponds to the natural year of registration).[8] For the date of birth, our survey data only provide information on the month and year of birth, while the lottery gives full dates (day/month/year) to determine the cutoffs that distributed conscripts to different military areas. For individuals born in the months without any cutoff, we can unambiguously determine the military destination. For the remaining individuals, we consider the set of all possible destinations.

Our final working sample has information on 1,093 individuals, of which 555 are men and 538 are women.[9] Note that we have applied the same sample restrictions regardless of the gender of the individuals. Given that women were exempt from military duties, they are an ideal placebo for the entire analysis.

Military migration: Once we have gathered the military destinations of the individuals in our sample, we construct the migration status for each of them, which is our treatment variable. We consider that an individual was assigned to military service outside his region if at least one of his possible military destinations does not contain his region of birth.[10] Otherwise, we assume that the individual did not migrate to another region.

National identity: To measure the national identity of the respondent we rely on a Likert five-item scale question in which individuals are asked to rate whether they feel "Only Spanish" (one), "More Spanish than gentilic of the region" (two), "As Spanish as gentilic of the region" (three), "More gentilic of the region than Spanish" (four), or "Only gentilic of the region" (five). From this scale, we construct the dummy variable "Spanish identity." This variable takes the value of one if the individual responds that he feels at least as Spanish as gentilic of the region (answers one, two or three on the Likert 5-item scale); otherwise, the variable takes the value of zero.[11]

Political behavior: For political behavior, we use two variables available in our survey data: voter turnout and voting for a regionalist party. Individuals in the survey waves from 2013, 2015 and 2017 declared their turnout and vote for the 2011, 2015 and 2016 Spanish General Elections, respectively. We coded as regionalist the parties that ran only in their region, did not form coalitions with other national parties, and defined themselves either

[8] Alternatively, we could use the region of residence at the time of the survey. However, we avoid this because the results of the lottery can affect current residence.

[9] In 1987, 9.04% of the potential recruits were exempted from military service due to a military surplus. These individuals were exempted from the military service by using the same lottery system, which exceptionally, for this year only, included a range of dates that allocated individuals to such surplus. We drop these individuals (55 observations) from the database.

[10] For the regions with provinces in multiple military areas (Castille-La Mancha and Castille-León), we consider that an individual was assigned to military service outside his region if his province of birth is not contained in the military area that he was assigned to. In Table A.4 of the online Appendix, we show that our results do not change when we exclude individuals from these regions.

[11] The original scale does not capture Spanish identity intensity (Guinjoan and Rodon, 2015). We discuss this transformation in section B of the online Appendix.

as regionalist or peripheral nationalist. The list of regionalist and non-regionalist parties according to the previous criteria is available in Table A.5 of the online Appendix. In the same table, we also propose a weaker and a stronger condition for being in the regionalist category. Table A.6 of the online Appendix shows that all the results are qualitatively the same across different definitions of "regionalist party".

Descriptive statistics: Table 1 shows the descriptive statistics for our sample and the two subsamples of individuals from the SSI and WSI regions. Approximately 37% of our sample is assigned to a military area outside their own region and is roughly the same for each type of region of origin.

Table 1: Descriptive Statistics

	Total		SSI		WSI	
	Mean	SD	Mean	SD	Mean	SD
% feel equal or more Spanish	83.6	37.1	90.8	28.9	60.3	49.1
% voted last general elections	83.2	37.4	84.0	36.7	80.9	39.4
% voted for regionalist party	5.6	23.0	0.5	6.9	22.1	41.7
% assigned to do MS outside region	37.3	43.1	37.1	42.9	37.9	44.0
Observations	555		424		131	

Notes: The sample, which is from the three ESGE waves, includes all male individuals born between September 1968 and December 1973. The first outcome is an indicator variable that takes the value of one if the individual feels that his Spanish identity is greater than or equal to his regional identity. The last outcome is an indicator variable that takes the value of one if the individual was assigned a nonzero probability of service outside his region.

3.2 Empirical specification

To estimate the effect of being assigned to military service in another region, we consider two main specifications. First,

$$y_{iprmv} = \alpha_0 + \alpha_1 \, \text{OUT}_{pmv} + \lambda_{rv} + \phi_m + \epsilon_{iprmv} \tag{1}$$

where y_{iprmv} is our outcome (e.g., identity or political behavior) for individual i born in province p of region r in month m in year v. The variable OUT_{pmv} is a dummy that takes the value of one if, according to our definition of military migration, an individual born in province p in month m and year v was allocated to serve in another region; otherwise, OUT_{pmv} takes the value of zero.[12] Since the yearly share of conscripts assigned to migrate varied between regions (not within), we include a set of dummies for region-by-year (λ_{rv}) fixed effects. We include a set of dummies for month of birth fixed effects (ϕ_m) to control

[12] In section C of the online Appendix we perform a sensitivity analysis of the results to different definitions of the treatment.

for the unobservable differences in births by calendar month.[13] Additionally, we include controls for whether parents were born in a region with a weak Spanish identity and fixed effects for the survey year.[14] Throughout the analysis, we cluster the standard errors (ϵ_{iprmv}) at the level of province and year of birth.

In a second specification, we allow for heterogeneity between conscripts from WSI regions and conscripts from SSI regions. Since a conscript from a SSI region is more likely to self-declare as Spanish and this identity presupposes a unitary notion of Spain, we expect that this conscript will consider other conscripts to be Spanish, regardless of their region of origin. On the contrary, a conscript from a WSI region is more likely to see conscripts from other regions differently from the conscripts that belong to his region. Accordingly, we expect to find a heterogeneous effect of intergroup contact for conscripts from WSI regions and SSI regions. Additionally, since conscripts from SSI regions are likely to have a higher baseline Spanish identity prior to service, there would always be less margin for any significant positive effect. This specification is

$$y_{iprmv} = \beta_0 + \beta_1 \,\mathrm{OUT}_{pmv} + \beta_2 \,\mathrm{Birth\,SSI}_r + \beta_3 \,\mathrm{OUT}_{pmv} \times \mathrm{Birth\,SSI}_r + \lambda_{rv} + \phi_m + \epsilon_{iprmv} \quad (2)$$

where $\mathrm{Birth\,SSI_r}$ is an indicator variable that takes the value of one if region r is of the SSI type. As mentioned before, we define the regions of Catalonia, Basque Country, Navarre and Galicia as WSI regions; we define all the rest as SSI regions.[15]

Our identification assumption is that after conditioning on region-by-year, serving in one's own or a different region is random. This is guaranteed by the structure and the timing of the lottery. Specifically, as we explained in the previous section, all decisions to defer or avoid the military service (such as pursuing academic studies) were made before the realization of the lottery. Additionally, being assigned to another region does not correlate with any pre-treatment observable characteristics (see Section A of the online Appendix).

The parameters of interest in our analysis are α_1 in equation (1) and β_1 and β_3 in equation (2). Parameter α_1 represents the change in the selected outcomes due to the contact induced by military service. Parameters β_1 and $(\beta_1 + \beta_3)$ represent the change in these outcomes for individuals from WSI regions and SSI regions, respectively. Note that our results should be interpreted as intention-to-treat estimates (ITTs) as we do not have information of the actual military service of the respondents.[16]

[13] Previous work finds that the quarter of birth affects the years of education (Angrist and Keueger, 1991) and other socioeconomic characteristics (Buckles and Hungerman, 2013).

[14] Our results are virtually unchanged if we do not include these controls.

[15] As a robustness check, we changed the WSI definition to exclude Galicia and/or Navarre, and the results are qualitatively the same. These estimates are reported in Table A.2 in the online Appendix.

[16] In Section D of the online Appendix we study the effect on identity for recruits without higher education for whom we anticipate an almost perfect compliance (above 90% according to our calculation).

4 Results

4.1 National identity

Table 2 presents the estimates for our Spanish identity outcome. Column (1) presents the estimate of α_1 in equation (1), while the rest of the columns report the estimates for the parameters of interest in equation (2) for different specifications. The estimate of α_1 in column (1) reveals that being assigned to military service in another region does not affect national identity. This lack of effect is explained by the heterogeneity reported in the remaining columns. We find that individuals from WSI regions are more likely to declare Spanish identity if they were assigned to military service in another region. The estimates for β_1 show an increase of approximately seventeen percentage points in the probability of expressing a Spanish identity; this increase is stable and consistent across these specifications. That is, being allocated to military service outside one's region of residence is associated with a 28% increase in Spanish identity among individuals from WSI regions with respect to the baseline mean.

In addition, for all specifications, our estimates of β_3 are not zero; that is, we find robust evidence for a heterogeneous effect of military service on national identity. Specifically, the estimated β_3 is negative with an absolute value similar to that of β_1, which suggests an impact close to zero among individuals from SSI regions. In fact, as reported at the bottom of Table 2, we fail to reject that the effect induced by the military service for individuals from SSI regions ($\beta_1+\beta_3$) is statistically different from zero. For the remainder of the paper, we use the specification in columns (4) and (5) with region-by-year and province-by-year fixed effects, respectively, as the preferred ones.

In Table 3, we further investigate the potential heterogeneous effects in different subsamples. In columns (1) and (2) we show that when we restrict the analysis to individuals from WSI regions, although we lose precision, the magnitude is very close to the one reported in the previous table. In columns (3) and (4), we increase our sample with individuals who resided in a different region at the time of the survey. Since we do not have direct information about the place of residence at the time of conscription, this new estimate is subject to a degree of measurement error. This error should be minimized for the sample of individuals who reside in their region of birth (those in our working sample). Indeed, our estimates are very similar to the estimates from Table 2 but with a slightly smaller point estimate. The overall conclusion across the different samples is the same: although the contact induced by military service increases Spanish identity, the impact is not significant for individuals from SSI regions.

Finally, in columns (5) and (6) of Table 3, we show the results for our placebo exercise. Given that only the male population was subject to mandatory military service, women constitute a natural placebo for our analysis. As expected, when we use the female sample, we observe not only a non-statistically significant effect (albeit with large standard errors) but also magnitudes that are considerably smaller than those observed for the male sample. This evidence supports that our findings above are not driven by unobservable factors that

Table 2: The effect of being assigned to military service outside one's own region on Spanish identity

	(1)	(2)	(3)	(4)	(5)
Out (β_1,α_1)	0.011	0.144	0.163**	0.166*	0.170*
	(0.034)	(0.090)	(0.080)	(0.086)	(0.093)
Birth SSI (β_2)		0.437***			
		(0.095)			
Out x Birth SSI (β_3)		-0.175*	-0.209**	-0.221**	-0.237**
		(0.097)	(0.089)	(0.093)	(0.101)
Observations	555	555	555	536	489
Mean dep. WSI	0.596	0.596	0.596	0.599	0.617
$\beta_1 + \beta_3$		-0.031	-0.047	-0.054	-0.067
H_0:$\beta_1 + \beta_3 = 0$, p-value		0.342	0.203	0.130	0.099
FE month	Yes	Yes	Yes	Yes	Yes
FE year	Yes	Yes	Yes	Yes	Yes
FE region			Yes	Yes	
FE region X year				Yes	
FE province					Yes
FE prov. X year					Yes

Notes: The outcome is an indicator variable that takes the value of one if the individual feels that his Spanish identity is greater than or equal to his regional identity. Birth SSI is a dummy variable equal to one if the individual was born in a SSI region. Out is a dummy variable equal to one if the individual was assigned to military service outside his own region. All specifications include fixed effects for year and month of birth. Standard errors are clustered at the year of birth-province level. $***p < 0.01, **p < 0.05, and *p < 0.1$

might be confounded with the lottery outcomes and are not captured by our set of fixed effects.[17]

4.2 Electoral impact

The estimates above demonstrate that individuals who were assigned to military service outside their own region were more likely to express a Spanish national identity 25 years later. However, one may wonder whether this effect on identity sentiment manifests in changes in political behavior. In this section, we examine the effects of the lottery on voting behavior.

Table 4 reports the estimates of regressing equation (2) on these outcomes. Panel A presents the results for the male sample and panel B shows the placebo analysis for the female sample. In Columns (1) and (2), we observe that individuals from WSI regions assigned to military service outside their region are approximately thirteen to seventeen percentage points more likely to have voted in the Spanish general elections than those who were not. In columns (3) to (6), we observe that these individuals are fourteen to nineteen

[17] In Table A.8 of the online Appendix using a seemingly unrelated regressions framework we also find a statistically significant effect for males but not for females. However, we should note, that the difference in the OUT coefficient (β_1) between males and females is not statistically significant.

Table 3: Robustness and placebo test

	Only WSI		W/migrants		Women	
	(1)	(2)	(3)	(4)	(5)	(6)
Out (β_1)	0.170*	0.138	0.144*	0.166**	0.011	0.105
	(0.098)	(0.127)	(0.076)	(0.079)	(0.104)	(0.113)
Out x Birth SSI (β_3)			-0.194**	-0.229***	0.025	-0.064
			(0.081)	(0.086)	(0.112)	(0.126)
Observations	142	133	630	569	525	475
Mean dep. WSI	0.599	0.617	0.605	0.623	0.553	0.564
$\beta_1 + \beta_3$			-0.050	-0.062	0.035	0.042
H_0:$\beta_1 + \beta_3 = 0$, p-value			0.100	0.076	0.447	0.438
FE region	Yes		Yes		Yes	
FE region X year	Yes		Yes		Yes	
FE province		Yes		Yes		Yes
FE prov. X year		Yes		Yes		Yes

Notes: For each sub-sample, indicated at top of the column, two specifications are reported. The first one using region fixed effects, the second one, province fixed effects. The first sample keeps only individuals born in regions with a WSI. The second sample additionally includes those individuals who, at the time of the survey, resided in a region different from that of birth. Finally, the last sample performs a placebo test on women, who did not participate in the lottery. The outcome is an indicator variable that takes the value of one if the individual feels that his Spanish identity is greater than or equal to his regional identity. Birth SSI is a dummy variable equal to one if the individual was born in a region with an SSI. Out is a dummy variable equal to one if the individual was assigned to military service outside his own region. All specifications include FE for year and month of birth. Standard errors are clustered at the year of birth-province level. $***p < 0.01, **p < 0.05, and *p < 0.1$

percentage points less likely to have voted for a regionalist party. These results, which we interpret as an increase in civic engagement in national political affairs, are consistent with our previous findings on Spanish identity. Once again, we find no effects for female individuals (Panel B).

One potential concern is that our results on party voting (and identity) are affected by the recent rise of secessionism in Catalonia, in particular, by the aggravation of the political conflict following the celebration of the October 2017 Independence Referendum and the imprisonment of secessionist Catalan leaders. However, the last election used in the 2017 survey occurred in 2016 (before the most disruptive events). Furthermore, we repeat the main analysis on both voting and identity without using the survey year 2017 (Table A.7), and the results do not change qualitatively.

4.3 Potential mechanisms

The first mechanism that is thought of when analyzing identity formation with military data is the direct impact that military service can have on subsequent feelings of identity. Given the role that militaries play in the political history of nations and because militaries are usually invoked by governments as institutions of national pride, a nation's military is

Table 4: The effect of being assigned to military service outside one's own region on voting behavior in the most recent general elections

	Voted		Voted Regionalist			
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (A): Males						
Out (β_1)	0.173**	0.126*	-0.119	-0.174*	-0.134*	-0.186**
	(0.071)	(0.074)	(0.080)	(0.091)	(0.078)	(0.087)
Out x Birth SSI (β_3)	-0.156*	-0.118	0.101	0.152*	0.115	0.163*
	(0.086)	(0.090)	(0.081)	(0.092)	(0.079)	(0.088)
Voted					0.087***	0.092***
					(0.025)	(0.032)
$\beta_1 + \beta_3$	0.017	0.009	-0.017	-0.022	-0.019	-0.023
H_0:$\beta_1 + \beta_3 = 0$, p-value	0.734	0.875	0.287	0.225	0.251	0.204
Panel (B): Females						
Out (β_1)	0.015	0.006	-0.022	-0.074	-0.023	-0.074
	(0.052)	(0.062)	(0.078)	(0.089)	(0.077)	(0.089)
Out x Birth SSI (β_3)	-0.076	-0.012	0.001	0.051	0.006	0.052
	(0.066)	(0.077)	(0.072)	(0.084)	(0.071)	(0.083)
Voted					0.067***	0.035
					(0.025)	(0.022)
$\beta_1 + \beta_3$	-0.061	-0.006	-0.020	-0.023	-0.016	-0.022
H_0:$\beta_1 + \beta_3 = 0$, p-value	0.184	0.907	0.116	0.091	0.209	0.095
FE region	Yes		Yes		Yes	
FE region X year	Yes		Yes		Yes	
FE province		Yes		Yes		Yes
FE prov. X year		Yes		Yes		Yes
Observations	525	475	525	475	525	475

Notes: Panel A (B) presents the analysis for the sample of men (women). For each outcome, indicated at the top of the column, two specifications are reported. The first one using region fixed effects, the second one, province fixed effects. The outcome for columns 1 and 2 is an indicator variable that takes the value of one if the individual voted, while the outcome for columns 3 to 6 is an indicator variable that takes the value of one if the individual voted for a regionalist party. Null and nonvoters are included as zero. Birth SSI is a dummy variable equal to one if the individual was born in a region with an SSI. Out is a dummy variable equal to one if the individual was assigned to military service outside his own region. All specifications include fixed effects for year and month of birth. Standard errors are clustered at the year of birth-province level. $***p < 0.01, **p < 0.05, and *p < 0.1$

hardly a neutral institution. Therefore, it is reasonable to expect that military conscription affects an individual's national identity either through engagement or indoctrination mechanisms. This mechanism would be consistent with the findings of some of the above-cited literature, which has explored the effects of military conscription on other aspects of an individual's mindset (Navajas et al., 2020; Fize and Louis-Sidois, 2020). However, although we do not deny that military service itself can affect an individual's national identity, the source of variation that we exploit rules out this effect as a potential interpretation for our findings. Our estimates come from exploiting the difference between movers and stayers from different regions, where both types are equally likely to serve in the military.

The mechanism behind our findings involves the random displacement of some individuals to another part of the country, where they necessarily will spend some time interacting and working with people from other regions.[18] In such circumstances, exposure to individuals from other regions may influence movers' feelings of identity, thereby strengthening the sense of a shared national identity.[19] This is the explanation that we have brought into contention throughout the text inasmuch as we believe that this explanation is the one most consistent with the source of variation that we exploit. To further explore this hypothesis, we study the effects of being assigned to another region on friendship formation. Having friends from another region is an indicator that conscripts had contact with colleagues and civilians from other regions.

Table 5 reports the estimates for equations (1) and (2) in Columns (1) and (2), respectively; these equations use an "outside friendship" dummy as a dependent variable, which takes the value of one if the individual declared that at least one person among his five closest friends or partner was born in a different region; otherwise, the variable takes the value of zero. According to the estimates in column (1), being allocated to military service outside the region of residence increases the probability that an individual reports at least one person in his inner circle who was born in a region other than his own by sixteen percentage points. This result is equivalent to an increase of approximately fifty-three percent with respect to the sample mean.

In column (2), we do not find evidence of heterogeneity between conscripts coming from WSI and SSI regions. We are careful at interpreting this result given the larger standard errors due to the reduction of the sample (the outcome is only available in the 2013 wave). Nevertheless, we believe that the possible absence of heterogeneity in friendship formation is compatible with the heterogeneity in national identity: having new friends from other regions could trigger a change in national identity only for those who did not share a common identity with them (i.e. conscripts from WSI regions).

[18] One may consider that the displacement of individuals to another region changes their exposure not only to individuals from other groups but also to other parts of the environment (e.g., institutions). We can consider this possibility by interpreting intergroup contact in a broader sense, specifically, one that includes exposure to any element of the outside group.

[19] Importantly, for most of the individuals the allocation takes place at the age of eighteen and this displacement is likely to be the first time outside their own region, which increases the importance of any exposure.

Table 5: Outside friendship

	Males		Females	
	(1)	(2)	(3)	(4)
Out (β_1,α_1)	0.163**	0.087	-0.003	-0.113
	(0.073)	(0.148)	(0.067)	(0.088)
Out x Birth SSI (β_3)		0.100		0.163
		(0.173)		(0.124)
Observations	201	201	205	205
Mean dep.	0.303	0.303	0.356	0.356
$\beta_1 + \beta_3$		0.187		0.049
H_0:$\beta_1 + \beta_3 = 0$, p-value		0.030		0.579

Notes: The dependent variable is a dummy equal to one if at least one person in the inner circle of individual i was born in a different region. Birth SSI is a dummy variable equal to one if the individual was born in a region with an SSI. Out is a dummy variable equal to one if the individual was assigned to military service outside his own region. Recall that the number of observations in this table is lower because information about friendship was available only in the 2013 wave. All specifications include fixed effects for year and region of birth. Standard errors are clustered at the year of birth-province level. $***p < 0.01, **p < 0.05, and *p < 0.1$

5 Concluding remarks

In this paper, we offer new evidence on how intergroup contact can contribute to nation building. We exploit the rules for the regional allocation of conscripts in the Spanish military service as a large-scale quasi-natural experiment. We find that, compared to their counterparts who stayed in their region, individuals from regions with a weak Spanish identity who were assigned to military service in another region were, many years later, more likely to identify themselves as Spanish. These individuals were also more likely to participate in national elections and less likely to vote for a regionalist party. Our source of variation is consistent with the hypothesis that intergroup contact helps shape the national identity of conscripts, which is supported by further evidence on friendship formation.

The findings in this study come in the context of calls for more research on how policies may be used to build collective identities and address legacies of regional and ethnic divisions. Although our study is a step in this direction, more research on how intergroup contact can be efficiently exploited in large-scale settings is needed. We believe that similar long-run effects can be achieved through other large-scale policies that also yield contact among different groups; examples of such policies include educational exchanges, national social services or measures that foster internal labor mobility. In this regard, the acclaimed success of recent experiences, such as the Erasmus program–the popular European Union

student exchange program–can be achieved through policies addressed to broader audiences, specifically in contexts with persistent regional and ethnic divisions.

References

Alesina, Alberto, Bryony Reich and Alessandro Riboni. 2020. "Nation-building, nationalism and wars." *forthcoming at Journal of Economic Growth* . 3

Alesina, Alberto, Johann Harnoss and Hillel Rapoport. 2016. "Birthplace diversity and economic prosperity,." *Journal of Economic Growth* 21:101—-138. 2

Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public goods and ethnic divisions." *The Quarterly Journal of Economics* 114(4):1243–1284. 2

Allport, Gordon W. 1954. *The Nature of Prejudice*. Cambridge, UK: Cambridge University Press. 2

Amat, Francesc. 2012. "Party competition and preferences for inter-regional redistribution in Spain." *South European Society and Politics* 17(3):449–465. 5

Angrist, Joshua D. 1990. "Lifetime earnings and the Vietnam era draft lottery: evidence from social security administrative records." *The American Economic Review* pp. 313–336. 4

Angrist, Joshua D and Alan B Keueger. 1991. "Does compulsory school attendance affect schooling and earnings?" *The Quarterly Journal of Economics* 106(4):979–1014. 9

Arbatlı, Cemal Eren, Quamrul H Ashraf, Oded Galor and Marc Klemp. 2020. "Diversity and conflict." *Econometrica* 88(2):727–797. 2

Aspachs-Bracons, Oriol, Irma Clots-Figueras, Joan Costa-Font and Paolo Masella. 2008. "Compulsory language educational policies and identity formation." *Journal of the European Economic Association* 6(2-3):434–444. 3

Bagues, Manuel and Christopher Roth. 2020. "Interregional contact and national identity.". 4

Bandiera, Oriana, Myra Mohnen, Imran Rasul and Martina Viarengo. 2018. "Nation-building Through Compulsory Schooling during the Age of Mass Migration." *The Economic Journal* 129(n/a):62–109. 3

Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg and Maisy Wong. 2019. "Unity in diversity? How intergroup contact can foster nation building." *The American Economic Review* 109(11):3978–4025. 3

Blumer, Herbert. 1958. "Race prejudice as a sense of group position." *Pacific sociological review* 1(1):3–7. 2

BOE. 1987. "Orden 423/38776/1987, de 9 de septiembre, por la que se dan normas para el sorteo de los mozos pertenecientes al reemplazo de 1988 y agregados al mismo.". https://www.boe.es/diario_boe/txt.php?id=BOE-A-1987-21828. 6

Boisjoly, Johanne, Greg J Duncan, Michael Kremer, Dan M Levy and Jacque Eccles. 2006. "Empathy or antipathy? The impact of diversity." *The American Economic Review* 96(5):1890–1905. 3

Buckles, Kasey S and Daniel M Hungerman. 2013. "Season of birth and later outcomes: Old questions, new answers." *Review of Economics and Statistics* 95(3):711–724. 9

Cáceres-Delpiano, Julio. 2019. "The Impact of Mandatory Military Service. Evidence from Spain." *UC3M working paper series* (2019-03). 4

Card, David and Ana Rute Cardoso. 2012. "Can compulsory military service raise civilian wages? Evidence from the peacetime draft in Portugal." *American Economic Journal: Applied Economics* 4(4):57–93. 4

Carrell, Scott E, Mark Hoekstra and James E West. 2019. "The impact of college diversity on behavior toward minorities." *American Economic Journal: Economic Policy* 11(4):159–82. 3

Casas, Agustin, Federico Curci and Antoni-Italo De Moragas. 2020. "Checks and Balances and Nation Building: The Spanish Constitutional Court and Catalonia." Mimeo. 4

Dahl, Gordon, Andreas Kotsadam and Dan-Olof Rooth. 2020. "Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams." *The Quarterly Journal of Economics* . 3

Dehdari, Sirus H and Kai Gehring. 2019. "The origins of common identity: Evidence from Alsace-Lorraine." Mimeo. 4

Dell, Melissa and Pablo Querubin. 2017. "Nation building through foreign intervention: Evidence from discontinuities in military strategies." *The Quarterly Journal of Economics* 133(2):701–764. 4

Depetris-Chauvin, Emilio, Ruben Durante and Filipe R Campante. 2020. "Building nations through shared experiences: Evidence from African football." *The American Economic Review* 110(5):1572–1602. 4

Desmet, Klaus, Joseph Flavian Gomes and Ignacio Ortuño-Ortín. 2020. "The geography of linguistic diversity and the provision of public goods." *Journal of Development Economics* 143:102384. 2

Easterly, William and Ross Levine. 1997. "Africa's growth tragedy: policies and ethnic divisions." *The Quarterly Journal of Economics* 112(4):1203–1250. 2

Erikson, Robert S. and Laura Stoker. 2011. "Caught in the Draft: The Effects of Vietnam Draft Lottery Status on Political Attitudes." *The American Political Science Review* 105(2):221–237. 4

Fize, Ettiene and Charles Louis-Sidois. 2020. "Military Service and Political Behavior: Evidence from France." *European Economic Review* 122. 4, 14

Galiani, Sebastian, Martín A Rossi and Ernesto Schargrodsky. 2011. "Conscription and crime: Evidence from the Argentine draft lottery." *American Economic Journal: Applied Economics* 3(2):119–36. 4

Guinjoan, Marc and Toni Rodon. 2015. "A scrutiny of the Linz-Moreno question." *Publius: The Journal of Federalism* 46(1):128–142. 7

Kukic, Leonard. 2019. "The last Yugoslavs: ethnic diversity, national identity and civil war." *Mimeo* . 3

La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Shleifer and Robert Vishny. 1999. "The quality of government." *The Journal of Law, Economics, and Organization* 15(1):222–279. 2

LaVanguardia. 1987. "Más de 260.000 mozos correspondientes al reemplazo de 1988 fueron protagonistas del primer sorteo único del servicio militar.".
URL: *http://hemeroteca.lavanguardia.com/preview/1987/11/16/pagina-18/32998166/pdf.html?search=mozos* 5

Lee, Neil. 2015. "Migrant and ethnic diversity, cities and innovation: Firm effects or city effects?" *Journal of Economic Geography* 15(4):769—-796. 2

Merlino, Luca Paolo, Max Friedrich Steinhardt and Liam Wren-Lewis. 2019. "More than just friends? School peers and adult interracial relationships." *Journal of Labor Economics* 37(3):663–713. 3

Michalopoulos, Stelios and Elias Papaioannou. 2016. "The long-run effects of the scramble for Africa." *American Economic Review* 106(7):1802–48. 2

Miguel, Edward. 2004. "Tribe or nation? Nation building and public goods in Kenya versus Tanzania." *World Politics* 56(3):327–362. 2

Montalvo, Jose G and Marta Reynal-Querol. 2005. "Ethnic polarization, potential conflict and civil war." *The American Economic Review* 95(3):796–816. 2

Montalvo, Jose G and Marta Reynal-Querol. 2021. "Ethnic Diversity and Growth: Revisiting the Evidence." *The Review of Economics and Statistics* 103(3):1–12. 2

Moreno, Luis. 2002. "Decentralization in Spain." *Regional Studies* 36(4):399–408. 4

Muñoz, Jordi. 2009. "From National-Catholicism to Democratic Patriotism? Democratization and reconstruction of national pride: the case of Spain (1981–2000)." *Ethnic and Racial Studies* 32(4):616–639. 4

Navajas, Gabriela Ertola, Paula A López Villalba, Martín A Rossi and Antonia Vazquez. 2020. "The Long-Term Effect of Military Conscription on Personality and Beliefs." *Review of Economics and Statistics* pp. 1–34. 4, 14

Okunogbe, Oyebola. 2018. *Does Exposure to Other Ethnic Regions Promote National Integration?: Evidence from Nigeria*. The World Bank. 3

Ottaviano, Gianmarco and Giovanni Peri. 2005. "Cities and Cultures." *Journal of Urban Economics* 58(2):304–337. 2

Ottaviano, Gianmarco and Giovanni Peri. 2006. "The economic value of cultural diversity: evidence from US cities." *Journal of Economic Geography* 6(1):9–44. 2

Pew Research Center. 2019. "Spring 2019 Global Attitudes Survey.". 3

Putnam, Robert D. 2007. "E pluribus unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize Lecture." *Scandinavian political studies* 30(2):137–174. 2

Samii, Cyrus. 2013. "Perils or Promise of Ethnic Integration? Evidence from a Hard Case in Burundi." *The American Political Science Review* 107(3):558–573. 4

Scacco, Alexandra and Shana S Warren. 2018. "Can social contact reduce prejudice and discrimination? Evidence from a field experiment in Nigeria." *The American Political Science Review* 112(3):654–677. 3