The Political Legacy of Violence: The Long-Term Impact of Stalin's Repression in Ukraine

Journal of Politics, forthcoming, 2017

Arturas Rozenas[1], Sebastian Schutte[2], and Yuri Zhukov[3]

[1]Department of Politics, New York University, Room 225, 19 W 4th St, New York, NY 10012, arturas.rozenas@nyu.edu
[2]Department of Politics and Public Administration & Zukunftskolleg, University of Konstanz, Room Y219, 78457 Konstanz, Germany, sebastian.schutte@uni-konstanz.de
[3]Department of Political Science, University of Michigan, 5700 Haven Hall, Ann Arbor, MI 48109-1045, zhukov@umich.edu

Political scientists have long been interested in how indiscriminate violence affects the behavior of its victims, yet most research has focused on short term military consequences rather than long-term political effects. We argue that large scale violence can have an intergenerational impact on political preferences. Communities more exposed to indiscriminate violence in the past will – in the future – oppose political forces they associate with the perpetrators of that violence. We document evidence for this claim with archival data on Soviet state violence in western Ukraine, where Stalin's security services suppressed a nationalist insurgency by deporting over 250,000 people to Siberia. Using two causal identification strategies, we show that communities subjected to a greater intensity of deportation in the 1940's are now significantly less likely to vote for 'pro-Russian' parties. These findings show that indiscriminate violence systematically reduces long-term political support for the perpetrator.

Keywords: indiscriminate violence, elections, civil war, Ukraine, causal inference[1]

[1]Replication data available at JOP's dataverse: http://tinyurl.com/h79gsj7. Arturas Rozenas is grateful for financial support from Hoover Institution, Stanford University. Sebastian Schutte is grateful for financial support by the EU FP7 Marie Curie Zukunftskolleg Incoming Fellowship Program (Grant #291784).

Can exposure to large scale violence shape the the political behavior of affected communities generations later? Previous research has highlighted the formative role of violence in political and economic development (Balcells, 2012; Blattman, 2009; Lupu and Peisakhin, 2015). Yet we know little about how violence affects *long-term* political preferences and attitudes, and whether this effect – previously studied at the individual level – also extends to communities.

We argue that indiscriminate violence reduces long-term political support for the perpetrator. Communities that experienced indiscriminate violence in the past tend to vote – generations later – against political forces they hold responsible for that violence. We empirically evaluate this claim with archival data on the Soviet campaign against nationalist rebels in western Ukraine (1943-1950) – a 'hard test' for our theoretical expectations. During this conflict, Stalin's security services deported over 250,000 local residents to Siberia, defeating the insurgency by physically removing its local support base. If even such militarily 'successful' uses of indiscriminate violence reduce political support in the long run, we can expect similar patterns to hold in cases where perpetrators were unable to militarily defeat their opponents at the time.

We employ two independent research designs to estimate the persistent effect of Soviet violence in western Ukraine on election results from 2004 to 2014. First, we use access to Soviet railroads as an instrumental variable, to exploit exogenous variation in deportations due to logistics. Second, we employ a fuzzy regression discontinuity design, exploiting idiosyncratic variation in repression levels due to the discretion of local secret police officials and Communist party leaders. The results are similar across the two designs: contemporary electoral support for 'pro-Russian' parties is substantially weaker in communities that the Soviet government repressed more heavily.

Our findings contribute to several strands of research. First is the vast literature on the dynamics of civil conflict, which has overwhelmingly focused on the short-term military effects of violence (Kalyvas, 2006; Lyall, 2009). By demonstrating that violence can resonate across generations, and potentially influence election results decades later, we show that extant research has been overlooking a significant part of the story. Second, our findings inform a wider debate on whether experiences of violence consolidate or undermine the political identities of affected communities (Balcells, 2012; Campbell, 1998; Lupu and Peisakhin, 2015). Our results support the somewhat provocative thesis that national identity can form in response to harsh repression (Kotkin, 2014), in line with earlier studies on cohesion and nationalism (Gellner, 1983). Finally, our study answers a lingering 'so what' question in political science research on historical episodes of violence. As we show, violent events that happened many decades ago are important for political science today because they persistently shape political preferences and identities.

Past violence, future politics

Our study speaks to two related, but previously disconnected areas of inquiry. The first is international relations literature on violence in civil conflict, which has examined the effects of indiscriminate violence on the battlefield, but not on the ballot box. The second is comparative politics literature on violence and political behavior, which has mainly examined the short-term effects of violence, at an individual, rather than communal level.

Research on armed conflict has long confronted the puzzle of why political actors use indiscriminate violence[1], what effect it has on its intended targets and the dynamics of conflict (see for example Kalyvas, 2006; Kocher, Pepinsky and Kalyvas, 2011; Lyall, 2009; Metelits, 2010; Weinstein, 2007). Some see indiscriminate violence as an effective military instrument to suppress rebellion – in the short term and at the local level, at least – because it raises the costs of continued resistance (Douhet, 1921) and depletes an opponent's pool of recruits and resources (Byman, 2016; Lyall, 2009; Zhukov, 2015). Others argue that indiscriminate violence is either ineffective or counter-productive (Arreguìn-Toft, 2001; Condra and Shapiro, 2012; Kalyvas, 2006; Kocher, Pepinsky and Kalyvas, 2011; Schutte, 2016). Since indiscriminate violence targets victims based on collective markers rather than their individual behaviors (e.g., where one lives, not what one has done), such violence cannot incentivize people to act in line with the punisher's wishes because they expect to be punished irrespective of what they do (Schelling, 1966). Moreover, indiscriminate violence can encourage people to act *against* the punisher precisely to avoid exposure to violence (Kalyvas, 2006; Kocher, Pepinsky and Kalyvas, 2011).

Largely missing from this debate are the far reaching political implications of *community-level* experiences of violence. The logic of arguments on both sides is closely tied to wartime conditions, when the imminent threat of violence colors every political decision. We do not currently know whether short-term military gains from indiscriminate violence can translate into long-term political success, or whether casualties among bystanders make it more difficult to win civilian loyalties in the long run.

Instead of focusing on long-term community-level effects, existing scholarship on the historical legacy of violence has mainly studied *individuals* who have directly experienced violence, and compares them to individuals without such experiences. For instance, Blattman (2009) finds that ex-combatants in Uganda are more likely to vote in elections, and Balcells (2012) argues that individual experiences of violence in the Spanish Civil War affected the ideological preferences of survivors. Lupu and Peisakhin (2015) show that Crimean Tatars' family experience of Soviet violence affects their political attitudes three generations later.

Beyond the individual level, several recent studies have found that exposure to wartime violence may increase broader patterns of political participation and potentially affect electoral outcomes (De Luca, Verpoorten et al., 2015; Grosjean, 2014). Research on Israeli and Turkish politics has shown that terrorist attacks tend to increase support for parties that take a more hawkish stance toward the attackers (Berrebi and Klor, 2008; Getmansky and Zeitzoff, 2014). Like research on civil conflict, most of these studies have focused on shorter time scales. What little empirical evidence exists on long-term effects suggests that the political impact of violence decays over time (Costalli and Ruggeri, 2016).

Theoretical expectations

We expect indiscriminate violence to reduce long-term political support for the perpetrator. While agnostic to the simultaneously coercive and alienating effects of indiscriminate violence during conflict, we argue that only the alienating effects exceed a conflict's duration and solidify into community-level narratives of past victimization.

The starting point of our argument is that it is not necessary for a given individual to experience violence directly to be impacted by it. Violence against some members of the community might impact the beliefs and behaviors of others who were not direct victims. This distinction is particularly important in the case of indiscriminate violence, where perpetrators select targets based on collective criteria ('where one lives' not 'what one does') and every member of a community is a potential target. If the state deports, arrests or kills 20 percent of a community's residents, it would not be unreasonable to think that the remaining 80 percent who did not directly experience this violence were nonetheless affected by it in a politically meaningful way. Such experience of common victimization can facilitate group cohesion and increase national identification (see Gellner, 1983; Stein, 1976).

On a generational timescale, one of the consequences of community-level exposure to indiscriminate violence is a shared belief in a common enemy or threat. Because collective experiences drive collective identity, and collective identity drives political behavior (Eyerman, 2001; Shamir and Arian, 1999), this backlash is likely to have a more lasting effect on political preferences than coercive threats from a long-gone perpetrator. While the coercive effects of indiscriminate violence operate only during wartime, these inflammatory effects may last considerably longer.

- *Hypothesis*: Past exposure to indiscriminate violence by actor i reduces future local support for actor i.

We test this hypothesis by examining the effect of Soviet-era repression on contemporary voting in western Ukraine. Beyond its historical significance as the deadliest conflict in

Europe since the Second World War, the Soviet government's campaign against western Ukrainian nationalists in 1943-1955 has an attractive empirical property: it is a hard test for an adverse relationship between indiscriminate violence and future political behavior. Conflict scholars often cite western Ukraine as a historical example where indiscriminate violence 'worked', at least in the narrow military sense of reducing rebel attacks (Byman, 2016; Ucko, 2016; Zhukov, 2015). The history of the conflict also reveals several plausibly exogenous sources of variation in violence, which allow us to see if these military gains came at the expense of long-term political loyalty.

Soviet violence in western Ukraine

In late 1942, western Ukraine was in a state of nature (see Magosci, 1996, 625-637): neither Germans, who had just seized the territory from the Red Army, nor the Soviets who occupied it briefly prior to the German invasion, could credibly claim to control it. In this expanding security vacuum, the sole agents of the Soviet government were red partisan units whose primary objective was to disrupt German communications and logistics in the rear, kill local German collaborators and, eventually, to re-establish Soviet rule (Gogun, 2012; Statiev, 2014). The partisans' primary local competitor was the Organization of Ukrainian Nationalists (OUN), which originally formed as an activist group seeking an independent, mono-ethnic Ukrainian state. In direct response to the advancements of Soviet partisans, OUN established its own armed wing, the Ukrainian Insurgent Army (UPA).

The Red Army regained formal control over western Ukraine in 1944. Due to the presence of nationalist insurgents, however, Soviet authorities could not establish viable political control in the countryside, and especially in those places where red partisans did not establish an early foothold. What followed was the most protracted and deadly domestic conflict the Soviet Union ever faced. During its combat phase from 1943 to 1955, as conservative estimates suggest, 127,454 persons were killed and another 266,206 were forcibly resettled to other parts of the USSR.[2]

On paper, Moscow's objective was simple: 'all identified supporters [of OUN-UPA] in Ukraine are to be arrested with confiscation of property and sent to Chernogorsky special camp.'[3] In reality, the Soviets faced an *information problem*: identifying nationalist supporters was difficult and the OUN-UPA went to great lengths to deter the local population from cooperating with Soviet authorities. UPA supreme commander Roman Shukhevych summarized the organization's policy: '[we] should destroy all those who recognize Soviet authority. Not intimidate but destroy' (Statiev, 2010, p. 131). Given the OUN-UPA's uncompromising approach to security and experience in underground operations, in many places, the Soviets lacked basic information on residents' loyalties.

Unable to identify the true 'subversives,' Soviet authorities applied the principle of 'collective responsibility' and began deporting large segments of the local population to Siberia and the Far East. Unlike earlier deportations that forcibly relocated entire ethnic groups (e.g. Chechens, Crimean Tatars), Soviet deportations in Ukraine formally targeted only the families of those suspected of association with nationalists or whose absence could not be accounted for (Statiev, 2010, p. 173). Interpretations of what constituted a 'nationalist supporter' varied widely, often implicating people with no actual involvement or ties to the insurgency. In line with Stalin's 'five percent rule' – as long as five percent of victims are guilty, indiscriminate violence is justified (Gregory, 2009) – the NKVD sometimes deported entire villages. When failing to find the families identified as subversives, the NKVD would sometimes deport another family from the village as a replacement (Statiev, 2010, p. 175).

Moscow saw deportation not as a 'gentler' substitute for lethal violence, but as an escalation. Archival data reveal that deportation occurred in especially violent areas where the NKVD had difficulty identifying OUN-UPA supporters, and where other forms of violence proved insufficient (Zhukov, 2015). Deportation was also an escalation over more general forms of displacement (Balcells and Steele, 2016; Greenhill, 2011). Rather than simply creating an environment in which civilians faced strong incentives to flee, the Soviets went further by forcibly removing civilians from their homes. In explicitly targeting families and children rather than combatants, deportation was arguably the most indiscriminate form of repression available to the state.

Deportation had a significant negative effect on nationalist violence. All else equal, the average district saw a 38.9 percent decline in OUN-UPA attacks immediately after cases of deportation, compared to a 7 percent increase after conventional counterinsurgency (Zhukov, 2015). This negative effect has reinforced popular perceptions of this conflict as one where indiscriminate violence 'worked' (Byman, 2016; Ucko, 2016).

A challenge in estimating the *long-term* effect of deportation on contemporary political preferences is the endogeneity of Soviet violence to the perceived scale of local resistance. Soviet authorities used these methods in what they suspected were nationalist strongholds, and these traditionally 'anti-Russian' areas may still have ended up voting 'anti-Russian' today, even without Soviet repression. As we show in the next section, however, the historical context reveals two plausibly exogenous sources of variation in Soviet violence: military logistics and the discretion of local leadership.

The Soviets' deportation capacity was locally limited by their logistical ability to forcibly remove thousands of people from their homes and transport them to far-away regions of the USSR. Large-scale deportations were least costly in areas with direct access to Soviet railroad networks (Zhukov, 2016). The Soviets deported, on average, 102 more people from districts located within 2 miles of the railroad than from districts located farther away.[4] These railroads had a far greater impact on Soviet deportations than on rebel operations. Unlike red partisans in WWII, the OUN-UPA focused a relatively small share of their military activity on railroad sabotage, and – as a result – nationalist violence was not significantly higher in rail-accessible districts.[5]

The topology of this rail network predated the conflict by several decades – the western Ukrainian section of the Soviet rail system was a joint creation of Austria-Hungary, imperial Russia and interwar Poland. This was not a system originally developed to facilitate mass resettlement to Siberia. An expansion of the network began in 1953, after the conflict largely subsided. Given this history, we can reasonably expect access to mid-century Soviet railroads to be both predictive of deportation's scale, and orthogonal to contemporary politics (we document this in more detail in Appendix 6).

A second, more idiosyncratic source of variation was the zeal and ambition of local Soviet bosses, which varied discontinuously across the borders of administrative districts. This variation was a consequence of the deliberately localized nature of the anti-OUN campaign. To alleviate the information problem, Nikita Khrushchev – then First Secretary of the Communist Party of Ukraine (CP(b)U) – relied on local party branches and local NKVD officials to oversee the implementation of repression. A resolution by the Lviv regional party committee summarized this delegation of authority:

> It is the personal responsibility of secretaries of district committees of the CP(b)U, NKVD and NKGB chiefs, chiefs of garrisons to make extensive use of the forms and methods recommended by the CP(b)U and personally by N. Khrushchev before March 1, to completely eliminate the remnants of the gangs of Ukrainian-German nationalists and 'OUN underground.'[6]

While central authorities in Moscow dictated the general flavor of repression, its implementation was the responsibility of local political bosses and security services at the district (*rayon*) level – the local nucleus of party and administrative organization in the Soviet political system.[7] The heads of district-level party committees (*raykom*) and district-level secret police branches (NKVD and NKGB) ultimately decided whom to arrest, kill, and deport. As a consequence, repression varied significantly from district to district – in part due to the idiosyncratic preferences and beliefs of local bosses. In the mostly rural Rivne oblast, for instance, the Soviets resettled 611 individuals from Volodymyrets' district, but only 32 from neighboring Morochne (Zarichne) district – a territory of comparable size and population. Similarly, the Soviets resettled 810 people from Horodok district in Lviv oblast, but 'only' 116 from neighboring Yavoriv district. Our empirical strategy exploits this local variation, to more rigorously take stock of whether and how indiscriminate Soviet violence still affects the politics of this region today.

Data

To estimate the effect of Soviet violence on Ukrainian voting, we rely on declassified archival data on political violence and polling-station level contemporary election results. The administrative unit at which the NKVD organized its operations was the district, or *rayon*.[8] Because rayon boundaries changed over time, we obtained historical data on administrative boundaries from contemporaneous official Soviet directories and military maps (Main Topographic Directorate of the USSR General Staff; Presidium of Supreme Soviet of USSR, Information-Statistical Division 1941/1946/1954).

To measure the type and intensity of Soviet violence, we used archival event data first analyzed by Zhukov (2015). These data draw on a combination of declassified incident reports from central, regional and local organs of NKVD, Communist Party of Ukraine (KP(b)U),[9] and collections of OUN-UPA documents (Sokhan' and Potichnyj, 2002/2003). We used historical administrative boundaries to create rayon level measures of the intensity of Soviet repression (number of people deported) from 1943 to 1955 and indicators of red partisan control in 1942. We also calculated rayon level event counts for OUN-UPA violence, which we use as a covariate in our analyses.

We measure the scale of Soviet indiscriminate violence as the *absolute* number of people deported per rayon. We use absolute numbers rather than proportions for several reasons. First, the deportation quotas Soviet authorities sent to local officials – and the reports local officials sent back – used absolute rather than proportional numbers (Gregory, 2009). Second, more importantly, Ukraine's contemporary public debate on Soviet-era victimization privileges absolute numbers, even in cases – like the Crimean Tatars – where the proportion was 100 percent.[10] Lastly, we do not have reliable data on the denominator – neither the pre-

WWII 1931 Polish census nor the 1959 Soviet census offer reliable local population estimates for the mid-1940s.[11]

To measure contemporary pro-Russian political preferences (our dependent variable), we construct the variable *pro-Russian vote margin*, which we define as the difference between vote shares received by 'pro-Russian' and 'pro-Western/pro-nationalist' parties or candidates. To construct this variable, we use polling station-level election data from the Ukrainian Central Electoral Commission (UCEC) from 2004 to 2014. Using their geographic coordinates, we match the polling stations with their respective historical rayons (which do not always align with contemporary administrative boundaries) and then create a rayon-level average margin of pro-Russian votes.[12]

The labels 'pro-Russian' and 'pro-Western/pro-nationalist' are meant to capture, however imperfectly, two related cleavages in Ukrainian politics: geopolitical orientation and national identity. Although these two cleavages have shaped the Ukrainian party system since its inception (Birch, 2000), they became most vivid after the Orange Revolution in 2004 (Kuzio, 2010). The Party of Regions, its successor Opposition Block, and the Communist Party of Ukraine (successor of the Soviet Communist party) have consistently advocated closer economic and security relations with Russia, recognition of Russian as Ukraine's second official language, and downplayed the importance of Ukrainian national identity. On the other side, parties like the Petro Poroshenko Bloc, Batkivshchina, Our Ukraine, and Yulia Tymoshenko Bloc have favored closer economic, cultural and military association with the EU and NATO, and – together with nationalist parties like Svoboda and Right Sector – have appealed to Ukrainian national identity in their electoral campaigns.[13]

To ensure that the compared units are geographically proximate and similar on many dimensions, we restrict our sample to the western oblasts of Ukraine.[14] Ukraine is a notoriously polarized country, with very distinct regional patterns of voting, often related to differences in imperial legacies (Peisakhin, 2012). Voters in western Ukraine, which was historically part of the Austro-Hungarian empire and was incorporated into the Soviet Union only after 1939,

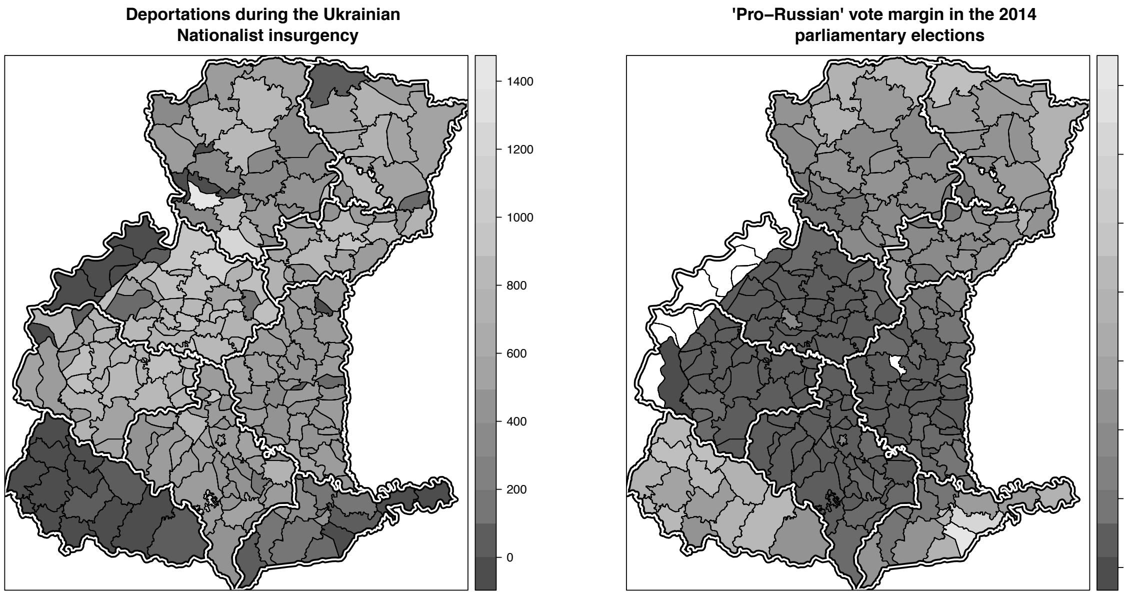


Figure 1: Historical violence and contemporary voting in western Ukraine. The figure on the left shows the counts of deported individuals. The right pane shows the 'pro-Russian' vote margin in the 2014 parliamentary elections. The westernmost rayons in white have no election data because the USSR returned them to Poland in 1945. Historical boundaries of *oblasts* appear in white. Please refer to Appendix 2 for residualized maps that account for systematic regional differences.

consistently vote more for Western-oriented and nationalist political candidates and parties. Conversely, voters in eastern Ukraine, incorporated into the Soviet Union by 1922, tend to vote for more 'pro-Russian' political forces. By restricting our sample to western Ukraine, we ensure that any variation in current voting patterns cannot be attributed to the historical legacies that make eastern and western Ukraine so different. We account for remaining intra-regional differences within western Ukraine – such as those between the historical regions of Volhynia, Galicia and Transcarpathia – through province (*oblast*)-level fixed effects.

Figure 1 shows the rayon-level distribution of Soviet-era deportations and 'pro-Russian' vote-margin in the 2014 parliamentary elections. Table 1 shows descriptive statistics for the main variables. Several patterns are worth noting. First, while a negative correlation between the maps is clearly visible, it is partially driven by historical legacy. Levels of repression and voting partially follow the 1918 partition of western Ukraine between Poland, Czechoslovakia, Hungary, and the USSR. For this reason, we use in all our estimations

Variable	Mean	Std. Dev.	Min.	Max.	N
Explanatory variables					
Deported individuals	521.85	289.13	0	1379	226
Soviet partisan control	0.18	0.39	0	1	226
Soviet partisan operations	0.12	0.32	0	1	226
Outcome variables					
'Pro-Russian' margin (parl. 2014)	-93.41	5.26	-98.41	-74.20	226
'Pro-Russian' margin (pres. 2014)	-92.98	6.44	-99.26	-65.21	226
'Pro-Russian' margin (parl. 2012)	-23.18	17.43	-48.26	45.33	226
'Pro-Russian' margin (pres. 2010)	-69.98	23.90	-95.83	41.32	226
'Pro-Russian' margin (parl. 2007)	-75.65	17.69	-94.68	-16.14	226
'Pro-Russian' margin (parl. 2006)	-62.52	16.75	-87.64	-1.51	226
'Pro-Russian' margin (pres. 2004)	-86.78	17.75	-99.42	-10.87	226
Instrumental variables					
Distance to railway (km)	7.04	11.17	0	70	226
Fraction forested	0.37	0.48	0	1	226
Covariates					
Intensity of rebel violence	25.93	25.16	0	199	226
Percent arable land	0.07	0.50	0	5	226
Days under German occupation	1083.90	80.69	953	1244	226
Urbanization	0.47	0.50	0	1	226
Russian-speaking in 1931 (%)	0.36	0.58	0	2.35	210

Table 1: Descriptive statistics.

fixed effects at the level of the oblast (white borders). Second, there is significant local variation *within* oblasts, with some spatially contiguous rayons seeing markedly different levels of repression and voting. This fact – clearly visible in residualized maps that adjust for regional differences (see Appendix 2) – helps us identify the impact of violence at the district level.

Results

To estimate the effect of WWII-era violence on present-day elections in Ukraine, we employ two identification strategies, at two levels of analysis. First, we analyze election outcomes over historical rayons using instrumental variables. Second, we analyze election outcomes at the polling-station level with a fuzzy regression discontinuity design.

Instrumental Variable Analysis

Our instrumental variable analysis builds on the following reduced-form equation:

$$y_{ij} = \alpha + \beta x_{ij} + \theta M_{ij} + \gamma L_j + \mathbf{v}_{ij} + \epsilon_{ij} \tag{1}$$

The dependent variable in Equation (1) is the 'pro-Russian' vote margin in rayon i, oblast j (y_{ij}).[15] The explanatory variable of interest is the legacy of violence (x_{ij}), as measured by the scale of Soviet deportation. M_i is a matrix of pre-treatment control variables, including the intensity of nationalist violence, the proportion of a rayon's territory that the Soviets designated as arable agricultural land, the number of days the rayon spent under German occupation, pre-WWII urbanization (dummy variable indicating the presence of at least one city or 'urban-type settlement'), and the share of the population that was Russian-speaking prior to the war, based on the 1931 Polish census.[16] L_j is a vector of oblast-level fixed effects, and ϵ_{ij} are iid errors. To account for spatial dependence, we include Moran eigenvectors $\mathbf{v}_{ij} = v_{ij1}, \ldots, v_{ijm}$ as synthetic covariates capturing residual autocorrelation (Dray, Legendre and Peres-Neto, 2006).[17]

The difficulty in identifying the effect of past violence on present-day voting is that contemporary political preferences may have deep historical roots, preceding even the Second World War. If prewar political loyalties drove some of the violence during and after WWII, then the β coefficient in Equation (1) may simply be capturing the effect of previous anti-Russian sentiment on today's anti-Russian sentiment.

To account for this possibility, we looked for an exogenous source of variation in Soviet deportations, which is likely to affect contemporary voting only through its influence on violent behavior. To this end, we used distance from each district to the railroad network as an instrumental variable for deportation. Because the structure of the rail network in

Ukraine was fixed and predated the bulk of the violence, this variable is causally prior to the Soviet counterinsurgency. To the extent that rail networks affect elections 70 years later, we can reasonably attribute that relationship to their effect on military activity in this period.

The railroad instrument would violate the exclusion restriction if – besides affecting military operations – rail networks drove Soviet-era economic development and industrialization and, in turn, economic policy preferences generations later. However, such patterns are more likely to attenuate than inflate the negative effect of indiscriminate violence on the 'pro-Russian' vote margin. Communities that saw more deportation due to their accessibility by rail (i.e. near built-up areas, cities, factories), should also have benefited more from subsequent Soviet economic development, which tended to favor urban development and industrialization at the expense of the countryside.

To facilitate the instrumental variable approach, we decompose the violence-voting relationship from Equation (1) into two stages. In the first stage, we use local railroad access as an instrument for violence:

$$x_{ij} = \mu + \zeta Z_{ij} + \phi M_{ij} + \kappa L_j + \mathbf{v}_{ij} + \mathbf{u}_{ij} + \upsilon_{ij} \tag{2}$$

where Z_{ij} is the instrumental variable, and $\mathbf{u}$ and $\mathbf{v}$ are sets of Moran eigenvectors for the first stage and second stage, respectively. The second-stage equation, which estimates the effect of violence on voting, is:

$$y_{ij} = \alpha + \beta^{(IV)} \hat{x}_{ij} + \theta M_{ij} + \gamma L_j + \mathbf{v}_{ij} + \epsilon_{ij} \tag{3}$$

where $\hat{x}_i$ are the fitted values of x_i from Equation (2), and the β^{IV} estimate is a local average treatment effect, representing the effect of deportation on voting in districts where deportation was more intense due the proximity of railroads.

We estimated Equations (2-3) for every Presidential and Parliamentary election between 2004 and 2014 for western Ukraine. Table 2 and Figure 2 report standardized coefficient estimates for $\hat{\beta}^{(IV)}$.[18] Appendix 3 reports the full set of parameter estimates.

Our first-stage estimates confirm that deportation was less intense in rayons less accessible by railroads. On average, a one-standard deviation increase in distance to the nearest railroad ($\uparrow$ 6.2 km) reduced the number of local number of deportees by .11 standard deviations ($\downarrow$ 31 deportees per district).[19]

	Dependent variable: 'pro-Russian' vote margin						
	2014 Parl.	2014 Pres.	2012 Parl.	2010 Pres.	2007 Parl.	2006 Parl.	2004 Pres.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Second stage							
Deportations ($\hat{\beta}^{(IV)}$)	-0.175^{**}	-0.098	-0.342^{***}	-0.301^{***}	-0.260^{***}	-0.246^{***}	-0.030
	(0.085)	(0.105)	(0.109)	(0.093)	(0.071)	(0.089)	(0.093)
OUN-UPA violence ($\hat{\theta}$)	0.052	0.004	0.074	0.094^{*}	0.068^{*}	0.093^{**}	-0.004
	(0.046)	(0.053)	(0.058)	(0.049)	(0.039)	(0.047)	(0.047)
First stage							
Distance to rail ($\hat{\zeta}$)	-0.106^{**}	-0.093^{*}	-0.102^{**}	-0.097^{**}	-0.107^{**}	-0.101^{**}	-0.091^{*}
	(0.050)	(0.047)	(0.049)	(0.048)	(0.049)	(0.048)	(0.046)
Covariates	Y	Y	Y	Y	Y	Y	Y
Oblast FE	Y	Y	Y	Y	Y	Y	Y
Moran eigenvectors	Y	Y	Y	Y	Y	Y	Y
Observations	217	217	207	217	215	216	218
Adjusted R^2	0.842	0.812	0.783	0.823	0.875	0.837	0.854
Weak instrument	5.082^{**}	4.495^{**}	4.926^{**}	5.676^{**}	10.485^{**}	7.221^{**}	5.749^{**}
Wu-Hausman test	1.419	0.003	5.169^{*}	4.397^{*}	5.323^{*}	3.475	0.95
Sargan test	15.473	10.918	10.604	5.455	13.899^{*}	8.568	8.423
Moran's I resid.	-2.728	-1.89	-2.698	-2.976	-3.347	-2.572	-1.709
Note:	*p<0.1; **p<0.05; ***p<0.01						

Table 2: Instrumental variable regression results. Standardized coefficients,with standard errors in parentheses. Intercept and control variables other than OUN-UPA violence not shown.

The second-stage results show that, in virtually every election cycle since 2004, the 'pro-Russian' vote margin was lower in rayons with a high level of post-WWII Soviet deportation. The effect is slightly stronger for parliamentary elections than for presidential ones, but the results are generally consistent in direction and magnitude.

In the parliamentary elections of 2014, the overall vote margin for the Opposition Bloc

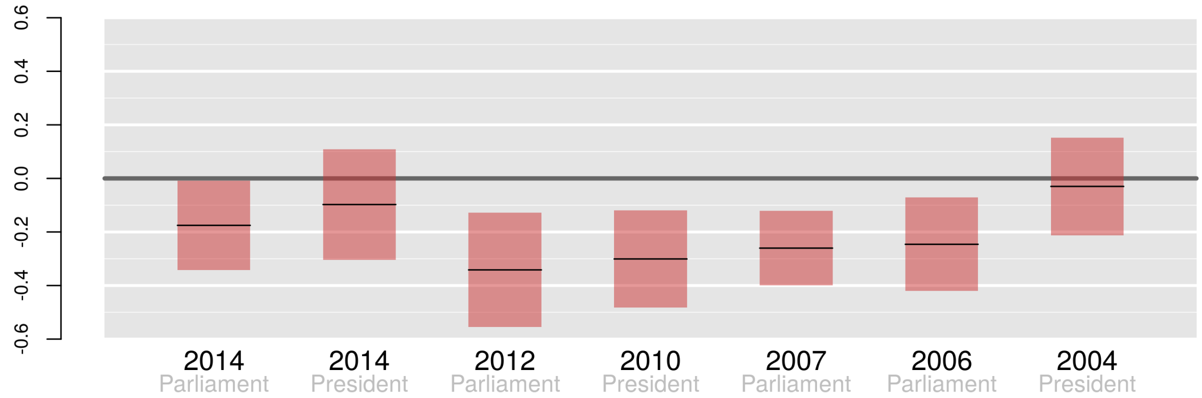


Figure 2: Standardized $\beta^{(IV)}$ coefficients: deportations. Quantities reported are standard deviation change in the 'pro-Russian' vote margin following a one-standard deviation increase in Soviet indiscriminate violence (deportations).

(successor to Party of Regions) was .18 standard deviation lower in areas of high Soviet deportation. These effects were even greater during the 2012 parliamentary elections – which the 'pro-Russian' parties won nationally. In those elections, the 'deportation effect' was a decline of .34 standard deviations in the vote margin, equivalent to a loss of 6 percent.

If we disaggregate the 'pro-Russian' bloc into individual parties, we see that the direction and size of the effects are consistent for mainstream 'pro-Russian' parties (e.g. Party of Regions, Opposition Bloc) but are even stronger for parties on the fringe. One such example is the Communist Party of Ukraine (CPU), successor to Ukraine's Soviet-era ruling party. Dissolved in 2015, following an investigation by the General Prosecutor for 'financing terrorism,' CPU had never been popular in western Ukraine. Yet there are places in this region where Communists under-performed their already low expectations – the same places the Soviets heavily repressed 70 years earlier. A standard deviation increase in Soviet deportation reduces CPU's vote share, on average, by .28 standard deviations, suggesting that Stalin's repression had a strongly negative effect on the electoral fortunes of the Soviets' ideological successors.

Past OUN-UPA violence, by comparison, appears to have had a weakly positive impact on the 'pro-Russian' vote. In most election cycles, however, areas hard-hit by OUN-UPA violence were neither more nor less likely to vote for 'pro-Russian' parties.

Regression Discontinuity Design

Stalin's repressive apparatus produced significant local variation, in part due to the autonomy it awarded to state and party agents implementing the deportations (Hagenloh, 2009, p. 9). As we discussed in the historical section, Moscow dictated the general policy line on western Ukraine, but the ultimate judgement on who was to be repressed depended on the idiosyncrasies of the local district-level heads of the NKVD and district-level committees of the Communist party (*raykom*). Consequently, some citizens were exposed to a larger risk of state violence simply by virtue of living in districts where local NKVD and Communist party officials were more willing to repress indiscriminately, more zealous to signal their loyalty to the central authorities through excessive repression, or simply more capable to repress.

As we show in Appendix 5.1, deportation levels varied drastically across spatially contiguous districts. For example, districts that experienced zero deportations were contiguous to other districts, which experienced anywhere from 0 to 1,400 deportations (with mean of 294). After adjusting for regional effects, spatial auto-correlation between districts explains only ten percent of variation in deportation levels. This numerical evidence suggests that idiosyncratic district-level factors were very important in determining patterns of Soviet repression.

Since the Soviets created district boundaries in western Ukraine only a few years prior to post-1943 deportations, we cannot attribute this cross-district variation to some underlying cross-district heterogeneity. Given the historical evidence, these differences are more likely driven by the idiosyncrasies of the security and party apparatus.

To explore the possibility that administrative boundaries drove geographic patterns of repression – and to provide further evidence of the long-term impact of Soviet deportations – we implement a fuzzy regression discontinuity design (Imbens and Lemieux, 2008). In this analysis, we measure the outcome variable ('pro-Russian' vote margin) at the precinct level, not at the district level as in the IV analyses. The idea here is that a precinct located in a historical district with more repressive local security service personnel and a more vigilant party apparatus experienced a greater risk of deportation than a nearby precinct located in a historically less repressive district. Accordingly, if 'pro-Russian' electoral support changes discontinuously as one moves across the border from a historically less repressive to a more repressive district, then we can plausibly attribute this discontinuous effect to the precinct's historical exposure to repression.

Let d_{ij} denote a forcing variable – the shortest-path distance (in kilometers) between polling station i in historical district j and the border to the nearest other historical district. d_{ij} takes negative values if district j had fewer deportations than the district in which precinct i is located. If district j had more deportations than i's district, then d_{ij} takes positive values. For example, the value $d_{ij} = -3$ means that precinct i (in district j) is located three kilometers away from the border of a historical district that had more deportations than district j; thus precinct i is located in an area with greater risk of deportations.[20]

For the estimates to be meaningful, we compare precincts located across the borders of districts with *contrasting* levels of historical deportations. In our main analysis, we only include those districts in which the levels of deportations were one standard deviation below or above the sample mean. In Appendix 5.3, we show that our results are robust to alternative contrast cut-off rules. In addition, to make sure we are comparing spatially proximate units, we only include polling stations located within ten kilometers of district borders. These restrictions result in a relatively small sample of 935 polling stations across all elections. To gain more statistical power, we pool data from all elections into a single analysis.

We start by estimating the following reduced-form equation:

$$\text{Pro-Russian Margin}_{ij} = \alpha_0 + \alpha_1 \mathbb{1}\{d_{ij} > 0\} + f(d_{ij}) + \epsilon_{ij}, \tag{4}$$

where $\mathbb{1}\{d_{ij} > 0\}$ is equal to one if the polling station is located in a historical district with

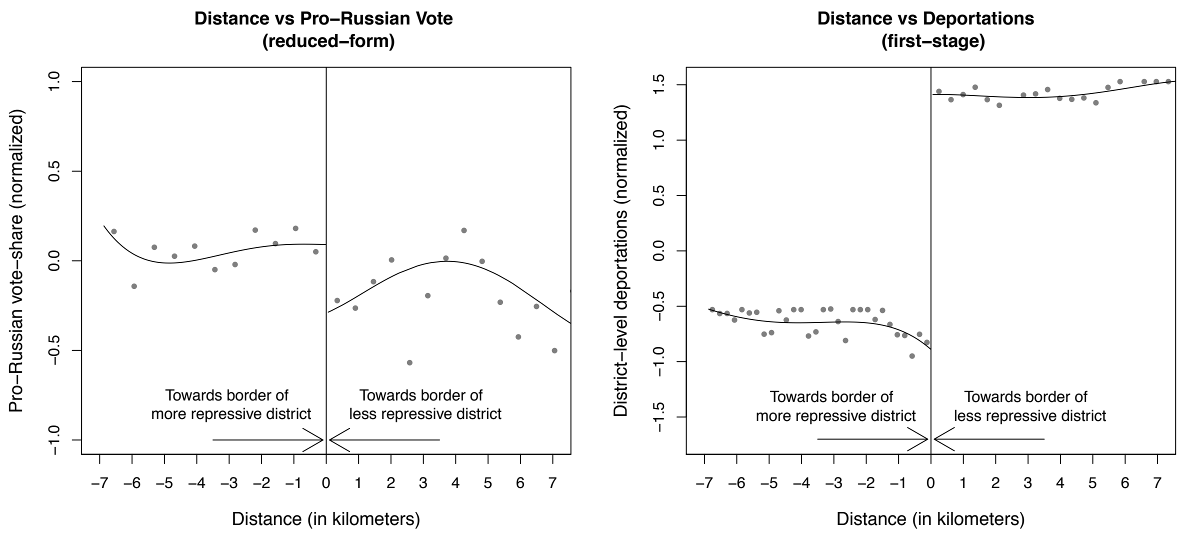


Figure 3: Reduced-form relationships between the instrument (distance from the contiguous district with more repression), deportations, and 'pro-Russian' vote.

a high level of repression and zero otherwise. $f(d_{ij})$ is a smooth function of distance to the nearest district border, which we estimate non-parametrically using the approach in Calonico, Cattaneo and Titiunik (2014). The parameter α_1 captures the spatial discontinuity effect of polling station i being located in a high-repression historical district on the contemporary 'pro-Russian' vote margin.

The estimated reduced-form relationship between the forcing variable and 'pro-Russian' support appears in the left panel of Figure 3, which uses the data-driven bin selection method (Calonico, Cattaneo and Titiunik, 2015). There is a clear discontinuous drop in 'pro-Russian' electoral support in more repressive districts. The estimated value of α_1 is equal to -0.39 (S.E. $= 0.17$, p-value < 0.03), meaning that 'pro-Russian' support drops by about 39 percent of a standard deviation as we move from precincts located in historically low-deportation districts to high-deportation districts.

The right panel of Figure 3 depicts the first-stage relationship between the forcing variable d_{ij} and standardized district-level deportations:

$$\text{Deportations}_j = a_0 + a_1 \mathbb{1}\{d_{ij} > 0\} + f(d_{ij}) + \epsilon_{ij}. \tag{5}$$

	Estimate	S.E.	p-value
1. Bias-correctd	-0.183	0.081	0.024
2. Bias-corrected with robust errors	-0.183	0.094	0.051
3. TSLS without variance weights	-0.171	0.081	0.036
4. TSLS with variance weights	-0.186	0.088	0.036

Table 3: Standardized Fuzzy RD estimates of the deportation effect on current pro-Russian vote. Standard errors in 3 and 4 are clusterred at the district level.

There is a very evident discontinuity between the forcing variable d_{ij} and deportation levels, which is not surprising given that we constructed the forcing variable to have this feature. Given the discontinuities in both reduced-form equations above, we can now estimate the second-stage regression:

$$\text{Pro-Russian Vote}_{ij} = \beta_0 + \beta_1 \text{Deportations}_j + g(d_{ij}) + e_{ij}, \tag{6}$$

where g is an unknown smooth function and β_1 represents the quantity of interest – the fuzzy RD effect of deportations on current 'pro-Russian' vote. Table 3 shows the estimated values of β_1 using four different estimation methods. For rows 1 and 2, we use bias-corrected local-polynomial estimation with conventional and robust standard errors, respectively, computed using the approach in Calonico, Cattaneo and Titiunik (2015). Row 3 shows the two-stage least squares (2SLS) estimate of β_I using quartic polynomial approximation to the functions f and g in Equations (5-6), following the procedure in Imbens and Lemieux (2008). Row 4 shows the weighted 2SLS estimate, which assigns greater weights to precincts located at the borders of districts with more contrasting levels of deportations.[21] Since our measure of deportations varies only at the district level, we cluster the standard errors by districts.

The fuzzy RD estimates are very similar across the four estimation methods: a one standard deviation increase in post-WWII deportations reduced support for 'pro-Russian' candidates from 17 to 19 percent of a standard deviation, depending on the estimation method. In three specifications (rows 1, 3, and 4), the estimated effect is significant at the 95 percent confidence level and in one specification (row 2) it is significant at the 90 percent confidence level. Importantly, the estimated effect of deportations is very similar in magnitude to the earlier IV estimates. The numeric similarity of the estimated effects across two independent designs (and using outcomes measured at different levels) is a strong indication these effects are not spurious.

Caveats

As in any non-experimental study, one should interpret the above results with a degree of caution, for several reasons.

First, the IV analyses invoked the exclusion restriction assumption, which requires that the instrument (distance to railways) affects the outcome ('pro-Russian' support) only through deportations and not some other channel unaccounted by the covariates (Angrist and Pischke, 2008). While we cannot entirely eliminate concerns related to the exclusion restriction, we conducted a placebo test for the railways instrument to see how well the exclusion restriction is likely to hold. In the test, we exploited the fact that Zakarpats'ka oblast in the south-western corner of Ukraine experienced very few deportations because the USSR annexed it only in 1946, after the main early waves of deportations. For this reason, we can treat Zakarpats'ka oblast as a placebo case: if the exclusion restriction is violated in a way that biases the $\hat{\beta}^{(IV)}$ coefficient in the negative direction, then distance to railroads should have a positive reduced-form effect on 'pro-Russian' support in precincts located in this placebo region. Our reasoning is that – since distance to railways was unlikely to drive deportations in Zakarpatiia – the reduced-form effect would indicate that distance to railways affected present political loyalties through some other channel than deportations. The results of this placebo test (Appendix 6) support the exclusion restriction assumption.

A second concern is that the railroad instrument may affect rebel violence as well as state violence (e.g. sabotage of government communications), and – because rebel violence may also affect the 'pro-Russian' vote – variation in OUN-UPA violence due to rail access may present an alternative pathway from railroads to political preferences.[22] To account for this possibility, we reran our analyses with OUN-UPA violence as the instrumented variable, and railroad access as the instrument. As we report in Appendix 4, there is little evidence that railroads drove variation in rebel attacks, or that rebel attacks drive voting today.

The fuzzy RD analysis also carries several important caveats. First, the settlements located across borders of districts with contrasting repression levels might not be similar in terms of their pre-treatment characteristics. In Appendix 5.2 we show that our design is well-balanced with respect to settlement-level religious and ethnic characteristics as measured by the 1921 Polish census.[23] It remains possible that the design is not well-balanced with respect to other characteristics for which we do not have settlement-level measures. Strong imbalance that would invalidate the design seems unlikely, because the Soviets created western Ukraine's administrative structure only mostly right before Germany invasion. It is doubtful that in such a short period of time – most of it under German occupation – spatially proximate settlements became very different from each other by virtue of belonging to different districts.

Second, there is the issue of compound treatment effects: if district bosses had discretion to repress based on their personal capabilities and political vigilance, they could have also done many other things differently (e.g., seize properties, distribute resources, mobilize local population). If so, we cannot ascribe the estimated fuzzy RD effect entirely to deportations. In the absence of reliable data on what the local NKVD and party bosses did in addition to deporting local populations, we cannot eliminate this concern systematically. However, we believe that, even if plausible, the bias created by the compound effect problem is not likely to be large: defeating the nationalist insurgency was by far the most important policy goal for both state security services and the Communist party apparatus during the time, and deportation was Moscow's main policy tool in this battle.

Finally, because they each rely on some assumptions that we cannot possibly test, the reported IV and fuzzy RDD estimates could potentially be more biased than a simple OLS estimate. Pooling data across all elections and adjusting for regional and election-level fixed effects, the OLS effect of deportations is equal to -0.061 (p-value < 0.01). In Appendix

7, we conduct the Caetano (2015) exogeneity test, which suggests that this OLS coefficient most likely *underestimates* the true effect. This is consistent with the fact that fixed effects OLS estimates are substantially smaller than IV and fuzzy RDD estimates.

Alternative Explanations

The results just presented indicate a strong and consistent relationship between the legacy of Soviet-era violence and voting patterns in contemporary western Ukraine. However, these empirical patterns are consistent with several explanations.

The explanation that we proposed, based on existing scholarship and our knowledge of the Ukrainian case, is that common experiences and memories of violence shape people's identities, and these identities shape the political preferences of communities impacted by violence (Eyerman, 2001; Shamir and Arian, 1999; **?**). Communities exposed to indiscriminate violence shared an experience in which every household was potentially at risk of repression, irrespective of their loyalty toward the Soviet regime. Such communities came to see the Soviet regime as targeting them less for their behavior, than for their identity as Galicians, Volhynians, or simply Ukrainians – an identity which local nationalists defined in opposition to Muscovite, Russian rule. The identity-forming effect of violence was further reinforced by the structural consequences of Soviet indiscriminate violence, which left a relatively high proportion of nationalist 'identity entrepreneurs' in affected communities.

Our proposed explanation rests on the claim that the indiscriminate nature of violence matters: because many Ukrainians fell prey to Soviet repression due to *who* they were, rather than their loyalty or disloyalty toward the Soviet state, deportations increased long-term anti-Russian sentiment. However, it is certainly plausible to argue that the nature of violence did not matter and that its long-term impact would have remained the same even if Soviet repression was highly discriminate.

While we cannot rule out this alternative explanation directly, we can assess its plausibility indirectly by comparing the effect of more indiscriminate Soviet deportations with the long-term political effect of more selective violence perpetrated by pro-Soviet partisans during World War II. Red partisans were state-sponsored guerrilla groups that formed and operated in Ukraine during Nazi occupation. In places where red partisans established an early foothold, they forged deeper links to local populations, and cultivated local informants. Here, the NKVD could later rely on existing networks of loyalists, who formed special 'extermination battalions' and local self-defense forces (Burds, 1997; Tkachenko, 2000), and used violence on a more discriminate basis (Statiev, 2014, p. 1545). Where the red partisans failed to establish an early foothold, the Soviets had little or no local loyalist network to fall back on after WWII, and compensated for a lack of local information with violence.[24] These patterns are consistent with other conflicts: violence is typically more extreme and indiscriminate where combatants lack territorial control, and are unable to distinguish opponents from bystanders (DeMeritt, 2015; Kalyvas, 2006).

If the nature of violence is indeed irrelevant, then territorial control by red partisans – which serves as a proxy measure for Soviet *capacity* for selective violence – should have a similar effect on current voting patterns as do deportations. We test this possibility by replicating the earlier IV analysis, with WWII-era partisan control as the treatment variable, and forest cover as an instrumental variable. Forest cover allowed the partisans to establish base camps behind German lines, by concealing their activities from surveillance. The forestation instrument is appropriately orthogonal to both the establishment OUN-UPA base camps (due to the nationalists' superior ability to hide in plain sight among the local population) and the extent of Soviet deportations.[25]

Figure 4 reports the results of this analysis (see Appendix 4 for full table of coefficients). The evidence strongly contradicts the notion that the nature of violence does not determine its long-term political impact. Rayons where red partisans established strongholds – and where Soviet authorities subsequently had better information – see significantly more 'pro-Russian' voting today. The normalized magnitude of this effect is one tenth of a standard

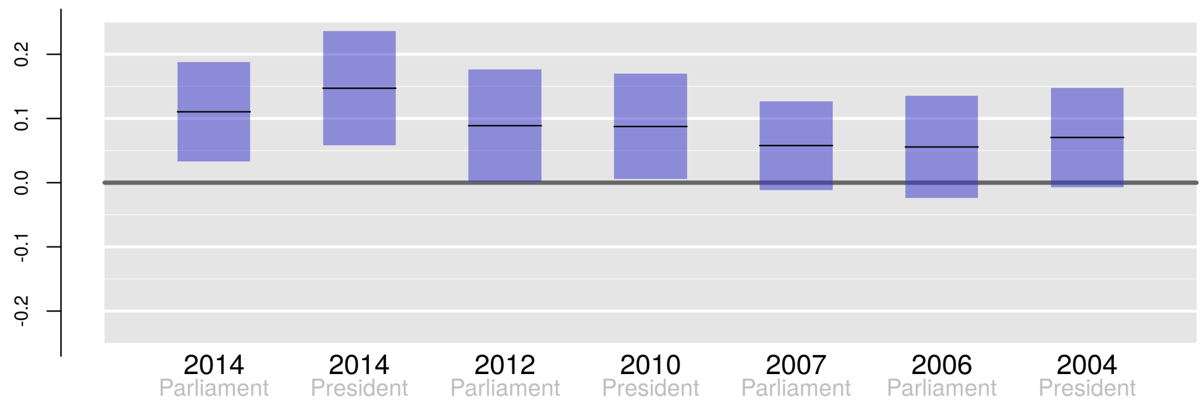


Figure 4: Standardized $\beta^{(IV)}$ coefficients: partisans. Quantities reported are standard deviation change in the 'pro-Russian' vote margin following a one-standard deviation increase in Soviet capacity for selective violence (partisan control in WWII).

deviation in the parliamentary elections of 2010-2014, and .15 standard deviations in the 2014 presidential race. While indirect, this test indicates that more selective tactics may have *improved* long-term support for Moscow.

The second alternative explanation is that the Soviet violence has significantly impacted ethno-linguistic composition. Hence, we see less support for 'pro-Russian' parties in more historically violent areas, because those areas simply have a lower proportion of Russian speakers. New settlers from eastern Ukraine and Russia may have viewed formerly repressed areas as hostile to Russian speakers, and instead settled in places they found more welcoming, like former partisan strongholds.

To test the plausibility of this alternative explanation, we implement two types of tests. First, we check whether Soviet violence had an impact on modern ethno-linguistic composition. Second, we check whether Soviet violence continues to affect voting after we adjust for modern ethno-linguistic composition.[26] If the effect attenuates significantly after this post-treatment adjustment, the 'ethnic composition mechanism' may be a plausible explanation of our results. To measure ethno-linguistic composition, we use rayon-level differences in the percentage of Russian to Ukrainian speakers, as measured in the 2001 Ukrainian census. The results, reported in Appendix 9, show that Soviet violence did not have a discernible impact on modern linguistic composition. Moreover, the estimated effect of Soviet violence on contemporary voting continues to be very similar to our earlier estimates, after controlling for post-treatment linguistic composition and other potential confounders.[27]

While the above analyses indicate that some alternative explanations of our findings are not plausible given the evidence, we certainly cannot rule out other possibilities, including the effect of repression on economic development, family structure, and social networks, among other things. Given the long-term political significance of trauma in post-conflict societies, understanding the intermediate effects of violence on the structure, identity and decision-making calculus of affected communities should be a priority area for future research.

Discussion

The main goal of this paper was to identify the long-term community-level political effect of historical violence. Using data on Soviet repression and contemporary voting in western Ukraine, we found that the Soviet policy of forced population resettlement had significant *negative* long-term effects on local support for 'pro-Russian' parties. We arrive at this result through two independent identification strategies – an instrumental variable design that exploits exogenous variation in violence due to railroads, and a regression discontinuity design that exploits variation due to local Soviet leadership. Both approaches yield the same finding: indiscriminate violence perpetrated by the Soviet state fomented strong anti-Soviet (and, consequently, anti-Russian) sentiment that persisted through generations.

Detecting such an effect should have been difficult in western Ukraine. Recent evidence suggests that Soviet authorities 'won' this conflict in part through their systematic use of deportation – physically uprooting the insurgency's local base of support. These population movements, and the decades of internal migration that followed, should have attenuated ties between Soviet repression and the communities that live in these places today. If the political effect of violence is consistently negative in Ukraine, then this effect is likely stronger in cases where the perpetrator initially 'lost.'

These findings have important implications for our understanding of how large scale violence shapes political identities and behaviors in the long run. First, our results suggest that the impact and trauma of violence extend beyond the individual level, and persist at the level of communities, reaching individuals who did not experience violence directly. Second, while we do find that violence potentially creates strong political preferences in opposition to the perpetrators of violence, this effect is not uniform and is highly dependent on the nature of violence.

More broadly, our results suggest that even militarily 'successful' cases of indiscriminate violence – of which western Ukraine was certainly an example at the time – may not prevent larger political losses in the long-term. The fact that 'anti-Russian' political preferences are currently strongest in localities where Soviet violence was most extreme shows that future scholars should look beyond short-term battlefield effects, and take a deeper look at how such violence shapes post-war political behavior.

Notes

[1]By 'indiscriminate violence' we do not mean only arbitrary or random violence, but also violence singling out particular groups, such as males of military age, members of a particular ethnic group, or inhabitants of localities that are under enemy control.

[2]GA SBU, F.13, D. 373, T. 103, L. 9-11.

[3]People's Commissariat of Internal Affairs (NKVD) order of January 7, 1944.

[4]The average district was 2 miles (3.2 km) from the closest railroad in 1945.

[5]The KS test statistic for the difference in means (28-21 rebel attacks) is insignificant, with $p = .19$.

[6]'Information of the Inspector of the Central Committee of the CPSU(b) N. Gusarov on the audit of Ukrainian party organization and the identification of a number of serious

shortcomings in its work,' July 13, 1946. Documents of 20th century http://doc20vek.ru/node/2329.

[7]Local discretion in the execution of violence is not unique to western Ukraine (Mitchell, 2004).

[8]Several *rayons* comprise an *oblast*, or province.

[9]Key archival sources include GARF, F. R-9478, Op. 1; GARF, F. R-9479, Op. 1; RGVA, F. 38650, Op. 1; TsDAGOU, F. 1, Op. 23.

[10]Consider, for example, the following headlines from Ukrainian media: 'How communists hastily killed more than 20,000 people in western Ukraine,' *TV Channel 24*, June 25, 2015; 'In one day more than 230,000 Crimean Tatars were deported,' *Uzhorod Inform*, May 27, 2016.

[11]One concern regarding the absolute measure is that population size may confound the relationship between deportation and political preferences. In Appendix 8, we show that this is unlikely.

[12]The UCEC provides the geographic coordinates for 2014 elections. For earlier elections, we geocoded the locations of polling stations using, in combination, the Yandex Maps service and GeoNames geographical database.

[13]Appendix 1 provides our exact coding scheme and classification of parties and candidates.

[14]These include Drogobych, Lviv, Rivne, Stanislaviv, Ternopil' and Volyn' oblasts (which the USSR annexed from Poland), as well as Chernivets'ka (from Romania) and Zakapats'ka oblasts (from Czechoslovakia). The resulting sample size comprises 226 administrative districts (rayons), based on the WWII-era Soviet administrative structure.

[15]The 'pro-Russian' vote margin is the percentage won by Party of Regions + Communist Party of Ukraine − all other parties.

[16]Because the Polish census does not include 16 rayons in Transcarpathia, we estimated models with the Russian language variable separately. Results are consistent with those reported below.

[17]The Moran eigenvector method diagonalizes the $N \times N$ connectivity matrix $\mathbf{C}$ (where $c_{ij} = 1$ if districts i and j share a border) to select the set of m eigenvectors with the largest achievable Moran's I coefficient of autocorrelation. To prevent multicollinearity, the algorithm extracts eigenfunctions of $[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{C}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']$, where $\mathbf{X} = [\iota_n \; X \; L]$ is the $N \times (k+1)$ matrix of covariates.

[18]To simplify comparisons across models, we report standardized coefficient estimates, which represent the number of standard deviations the outcome would change following a one-standard deviation increase in the explanatory variable.

[19]$\hat{\zeta}$ estimates vary slightly across elections due to different sets of Moran eigenvectors being included in first-stage regressions.

[20]Note that since we are using one-dimensional forcing variable, our design is not the standard *geographic* RDD in which a two-dimensional forcing variable is used (Keele and Titiunik, 2014). Using a two-dimensional forcing variable is complicated in our setting because we employ the fuzzy RD design. The main reason to use a two-dimensional forcing variable is to avoid comparing units that are distant from each other even though they are both close to a common border. This concern is quite negligible in our case because the average length of borders between districts is only 24.4 kilometers–units that are close to a border are also close to each other.

[21]Such weighting accounts for the fact that an RD design is most sensible in cases where neighboring districts have highly contrasting deportation levels.

[22]We are grateful to an anonymous reviewer for raising this point.

[23]For the RDD balance test, we need *settlement*-level demographic data, which we were able to obtain from 1921, but not from 1931 Polish census used in our IV analyses. As noted by Kopstein and Wittenberg (2011), who also use these census data, the 1921 census over-counts the Polish population. Since the variation in over-counting is unlikely to be correlated with *future* deportations, this systematic error in census data is unlikely to bias the results.

[24]For example, in the five districts of Stanislaviv oblast previously controlled by partisans, Soviet authorities conducted 40 percent fewer operations during the conflict than in other parts of the oblast. Similarly, they conducted 25 percent fewer operations in partisan-held districts of Rivne oblast.

[25]Districts with above-average nationalist activity were neither more nor less heavily forested that ones where they were less active. Districts the partisans controlled were 46 percent more heavily forested than the regional average.

[26]We use the same type of IV specification as above in equation 1, with modern ethno-linguistic composition as an additional covariate.

[27]In Appendix 8, we also discuss (and rule out) the possibility that our results are driven by the variation of pre-WWII and contemporary population/urbanization levels.

Acknowledgements

We thank Leonid Peisakhin, Christopher Sullivan, Jason Wittenberg, participants of the International Studies Association 2016 Annual Meeting in Atlanta, GA, and the Conference on Micro-Comparative Studies of 20th Century Conflicts, at Yale University. We also thank Jeffrey S. Kopstein and Jason Wittenberg for providing us with 1921 Polish census data. Finally, we thank Anastasia Vlasenko for excellent research assistance.

References

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Arreguìn-Toft, Ivan. 2001. "How the Weak Win Wars: A Theory of Asymmetric Conflict." *International Security* 26(1).

Balcells, Laia. 2012. "The Consequences of Victimization on Political Identities Evidence from Spain." *Politics & Society* 40(3):311–347.

Balcells, Laia and Abbey Steele. 2016. "Warfare, political identities, and displacement in Spain and Colombia." *Political Geography* 51:15–29.

Berrebi, Claude and Esteban Klor. 2008. "Are voters sensitive to terrorism? Direct evidence from the Israeli electorate." *American Political Science Review* 102(3):279–301.

Birch, Sarah. 2000. *Elections and democratization in Ukraine.* MacMillan.

Blattman, Christopher. 2009. "From violence to voting: War and political participation in Uganda." *American Political Science Review* 103(02):231–247.

Burds, Jeffrey. 1997. "AGENTURA: Soviet Informants' Networks & the Ukrainian Rebel Underground in Galicia, 1944-1948." *East European Politics and Societies* 11(1):89–130.

Byman, Daniel. 2016. "'Death Solves All Problems': The Authoritarian Model of Counterinsurgency." *Journal of Strategic Studies* 39(1):62–93.

Caetano, Carolina. 2015. "A test of exogeneity without instrumental variables in models with bunching." *Econometrica* 83(4):1581–1600.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6):2295–2326.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2015. "Optimal data-driven regression discontinuity plots." *Journal of the American Statistical Association* 14(4): 1753–1769.

Campbell, David. 1998. *National deconstruction: Violence, identity, and justice in Bosnia.* University of Minnesota Press.

Condra, Luke and Jacob Shapiro. 2012. "Who takes the blame? The strategic effects of collateral damage." *American Journal of Political Science* 56(1):167–187.

Costalli, Stefano and Andrea Ruggeri. 2016. "The long-term electoral legacies of civil war in young democracies: Evidence from Italy." Working Paper.

De Luca, Giacomo, Marijke Verpoorten et al. 2015. "Civil war and political participation: evidence from Uganda." *Economic Development and Cultural Change* 64(1):113–141.

DeMeritt, Jacqueline HR. 2015. "Delegating Death Military Intervention and Government Killing." *Journal of Conflict Resolution* 59(3):428–454.

Douhet, Giulio. 1921. *Command of The Air.* Air Force History and Museums Program.

Dray, Stéphane, Pierre Legendre and Pedro R. Peres-Neto. 2006. "Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM)." *Ecological Modelling* 196(3):483–493.

Eyerman, Ron. 2001. *Cultural trauma: Slavery and the formation of African American identity.* Cambridge University Press.

Garcıa-Ponce, Omar and Benjamin Pasquale. 2013. "How Political Violence Shapes Trust in the State." Working Paper.

Gellner, Ernest. 1983. *Nations and Nationalism.* Cornell University Press.

Getmansky, Anna and Thomas Zeitzoff. 2014. "Terrorism and voting: The effect of rocket threat on voting in Israeli elections." *American Political Science Review* 108(03):588–604.

Gogun, Aleksandr. 2012. *Stalinskie kommandos: Ukrainskie partizanskie formirovaniya, 1941-1944*. Moscow: ROSSPEN.

Greenhill, Kelly M. 2011. *Weapons of mass migration: forced displacement, coercion, and foreign policy*. Cornell University Press.

Gregory, Paul R. 2009. *Terror by quota: state security from Lenin to Stalin: an archival study*. Yale University Press.

Grosjean, Pauline. 2014. "Conflict and social and political preferences: Evidence from World War II and civil conflict in 35 European countries." *Comparative Economic Studies* 56(3):424–451.

Hagenloh, Paul. 2009. *Stalin's Police: Public Order and Mass Repression in the USSR, 1926-1941*. Washington, DC: Woodrow Wilson Center Press.

Imbens, Guido W and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of econometrics* 142(2):615–635.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge University Press.

Keele, Luke J and Rocio Titiunik. 2014. "Geographic boundaries as regression discontinuities." *Political Analysis* 23(1): 127–155.

Kibris, Arzu. 2011. "Funerals and elections: The effects of terrorism on voting behavior in Turkey." *Journal of Conflict Resolution* 55(2):220–247.

Kocher, Matthew A., Thomas B. Pepinsky and Stathis Kalyvas. 2011. "Aerial Bombing and Counterinsurgency in the Vietnam War." *American Journal of Political Science* 55(2):201–218.

Kopstein, Jeffrey S and Jason Wittenberg. 2011. "Deadly communities: Local political milieus and the persecution of Jews in occupied Poland." *Comparative Political Studies* 44(3):259–283.

Kotkin, Stephen. 2014. "Stalin, Father of Ukraine?" *New York Times, Nov. 27*.

Kuzio, Taras. 2010. "Nationalism, identity and civil society in Ukraine: Understanding the Orange Revolution." *Communist and Post-Communist Studies* 43(3):285–296.

Lupu, Noam and Leonid Peisakhin. 2015. "The Legacy of Political Violence across Generations." Working Paper.

Lyall, Jason. 2009. "Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya." *Journal of Conflict Resolution* 53(3):331–362.

Magosci, Paul R. 1996. *A History of Ukraine*. University of Toronto Press.

Martinez, Carla and T. Clifton Morgan. 2011. "Deterring Rebellion." *Foreign Policy Analysis* 7:295–316.

Metelits, Claire. 2010. *Inside insurgency*. New York University Press.

Mitchell, Neil. 2004. *Agents of atrocity: Leaders, followers, and the violation of human rights in civil war*. Springer.

Peisakhin, Leonid. 2012. *In History's Shadow: Persistence of identities and contemporary political behavior*. PhD thesis. Yale University.

Presidium of Supreme Soviet of USSR, Information-Statistical Division, ed. 1941/1946/1954. *SSSR: Administrativno-territorial'noye delenie soyuznykh respublik*. Moscow: Izd. 'Izvestiya Sovetov Deputatov Trudyashchikhsya SSSR'.

Schelling, Thomas C. 1966. *Arms and Influence*. New Haven: Yale University Press.

Schutte, Sebastian. 2016. "Violence and Civilian Loyalties: Evidence from Afghanistan." *Journal of Conflict Resolution* (forthcoming).

Shamir, Michal and Asher Arian. 1999. "Collective identity and electoral competition in Israel." *American Political Science Review* 93(02):265–277.

Sokhan', P. and P. Potichnyj, eds. 2002/2003. *Litopys UPA, New Series.* Vol. 4-7 Kyiv/Toronto: State Committee for Archives of Ukraine, Central State Archive of Public Associations of Ukraine.

Statiev, Alexander. 2010. *The Soviet Counterinsurgency in the Western Borderlands.* New York: Cambridge University Press.

Statiev, Alexander. 2014. "Soviet Partisan Violence against Soviet Civilians: Targeting Their Own." *Europe-Asia Studies* 66(9):1525–1552.

Stein, Arthur A. 1976. "Conflict and cohesion a review of the literature." *Journal of conflict resolution* 20(1):143–172.

Tkachenko, Sergei. 2000. *Povstancheskaya armiya: taktika borby.* Minsk and Moscow: Harvest AST.

Ucko, David H. 2016. "'The People are Revolting': An Anatomy of Authoritarian Counterinsurgency." *Journal of Strategic Studies* 39(1):29–61.

Weinstein, Jeremy M. 2007. *Inside Rebellion.* Cambridge University Press.

Zhukov, Yuri M. 2015. "Population Resettlement in War: Theory and Evidence from Soviet Archives." *Journal of Conflict Resolution* 59(7):1155–1185.

Zhukov, Yuri M. 2016. *Economic Aspects of Genocide, Mass Killings, and Their Prevention.* Charles H. Anderton and Jürgen Brauer, eds. Oxford University Press chapter "On the Logistics of Violence."

Online Appendix

The Political Legacy of Violence: The Long-Term Impact of Stalin's Repression in Ukraine

Contents

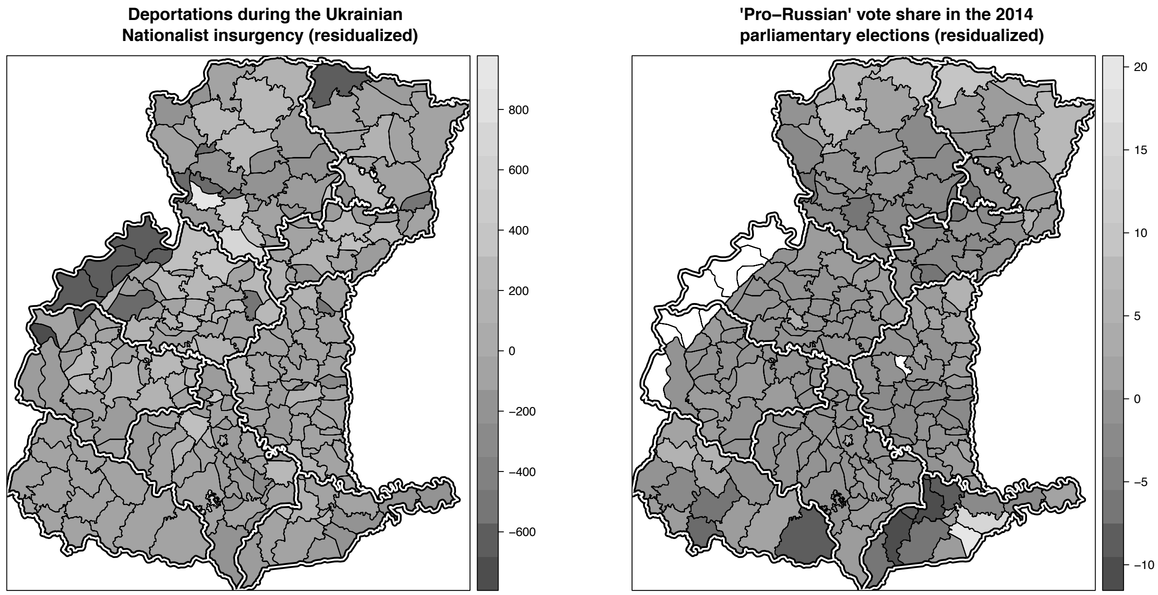
1 Classifying Political Parties and Candidates

We code political parties and candidates as pro-Western if their election manifestos and/or campaigns explicitly advocated for Ukraines membership in the European Union or NATO or promoted the strengthening of economic, social, or military ties with Europe. Otherwise, if they call for closer cooperation with the Customs Union of Russia, Kazakhstan, and Belarus or the Eurasian Economic Union are coded as pro-Russian. For presidential contenders in 2014 elections, we classify as 'pro-Western' those that served exclusively in the Viktor Yushchenko or Yulia Tymoshenko administrations or who were active on the side of the anti-Yanukovych protesters during the Euromaidan and those who served exclusively in the Yanukovych government are labeled as pro-Russian. In 2004 and 2010 presidential run-offs, only two candidates were competing. Table below shows the classification for all parties and candidates that received more than 5 and 3 percent of national vote, respectively.

Pro-Russian	**Pro-Wester/pro-nationalist**
2004 December 26, Presidential election (re-run)	
Viktor Yanukovich	Viktor Yushchenko
2006 Parliamentary election	
Party of Regions, Communist Party	Yulia Tymoshenko Bloc, Our Ukraine
2007 Parliamentary election	
Party of Regions, Communist Party	Yulia Tymoshenko Bloc, Our Ukraine
2010 Presidential election (run-off)	
Viktor Yanukovich	Yulia Tymoshenko
2012 Parliamentary election	
Party of Regions, Communist Party of Ukraine	All-Ukrainian Union (Batkivshchina), UDAR, Svoboda
2014 Presidential election	
Serhiy Tihipko, Mykhailo Dobkin	Petro Poroshenko, Yulia Tymoshenko, Oleh Lyashko, Anatoliy Hrytsenko
2014 Parliamentary election	
Opposition Bloc	Peoples Front, Petro Poroshenko Bloc, Self-Reliance Union, Radical Party of Oleh Liashko, All-Ukrainian Union Fatherland

2 Residualized Maps

Figure S1: **Historical violence and contemporary voting preferences.** The maps below show residuals from an *Oblast*-level Fixed Effects estimation that accounts for systematic regional differences. The figure on the left shows residuals for deported individuals. The right pane shows residuals for the 'pro-Russian' vote margin in the 2014 parliamentary elections. The boundaries of *Oblasts* are shown in green.



The maps displayed in the main text show raw values of 'pro-Russian' vote margins and historical deportations. While the negative correlation between the two is clearly visible, it is partially driven by historical legacy: after the defeat of the Ukrainian People's Republic in 1918, Western Ukraine was partitioned between Poland, Czechoslovakia, Hungary, and the Soviet Union based on the Treaty of Riga. Levels of Stalin-era repression and contemporary voting patterns follow this partition to some extent.

To account for this effect, the Instrumental Variable Design in the main text uses *Oblast*-level Fixed Effects. In Figure S1 above, we show the residuals from an *Oblast*-level Fixed Effects estimation. The negative correlation between historical repression and contemporary voting remains visible even as regional effects are taken into account.

3 Full Instrumental Variable (IV) Regression Results

Table S1 reports the full results of the models summarized in Figure 2 of the main text.

Table S1: **Instrumental variable regression results**. Standardized coefficients,with standard errors in parentheses. Coefficients for intercept and control variables shown.

	Dependent variable: 'pro-Russian' vote margin						
	2014 Parl.	2014 Pres.	2012 Parl.	2010 Pres.	2007 Parl.	2006 Parl.	2004 Pres.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Second stage							
Soviet deportations ($\hat{\beta}^{(IV)}$)	−0.175**	−0.098	−0.342***	−0.301***	−0.260***	−0.246***	−0.030
	(0.085)	(0.105)	(0.109)	(0.093)	(0.071)	(0.089)	(0.093)
Covariates ($\hat{\theta}$)							
OUN-UPA violence	0.052	0.004	0.074	0.094*	0.068*	0.093**	−0.004
	(0.046)	(0.053)	(0.058)	(0.049)	(0.039)	(0.047)	(0.047)
Agricultural land	−0.003	−0.008	−0.009	0.027	−0.003	0.002	0.002
	(0.028)	(0.030)	(0.033)	(0.029)	(0.025)	(0.028)	(0.027)
Duration of German occupation	−0.030	−0.058	0.030	−0.017	−0.095**	−0.059	−0.096*
	(0.053)	(0.057)	(0.063)	(0.056)	(0.048)	(0.054)	(0.051)
Urbanization	−0.024	−0.050	−0.006	0.026	0.019	0.021	0.009
	(0.034)	(0.037)	(0.040)	(0.036)	(0.030)	(0.034)	(0.033)
First stage							
Distance to rail ($\hat{\zeta}$)	−0.106**	−0.093*	−0.102**	−0.097**	−0.107**	−0.101**	−0.091*
	(0.050)	(0.047)	(0.049)	(0.048)	(0.049)	(0.048)	(0.046)
Covariates ($\hat{\phi}$)							
OUN-UPA violence	0.365***	0.369***	0.388***	0.368***	0.362***	0.364***	0.368***
	(0.049)	(0.050)	(0.051)	(0.049)	(0.048)	(0.049)	(0.050)
Agricultural land	0.023	0.023	0.023	0.023	0.026	0.026	0.023
	(0.042)	(0.043)	(0.044)	(0.042)	(0.041)	(0.042)	(0.042)
Duration of German occupation	0.041	0.042	0.053	0.042	0.036	0.038	0.043
	(0.080)	(0.081)	(0.082)	(0.080)	(0.078)	(0.080)	(0.080)
Urbanization	0.085*	0.085*	0.077	0.084*	0.087*	0.086*	0.084*
	(0.050)	(0.051)	(0.051)	(0.050)	(0.049)	(0.050)	(0.050)
Covariates	Y	Y	Y	Y	Y	Y	Y
Oblast FE	Y	Y	Y	Y	Y	Y	Y
Moran eigenvectors	Y	Y	Y	Y	Y	Y	Y
Observations	217	217	207	217	215	216	218
Adjusted R^2	0.842	0.812	0.783	0.823	0.875	0.837	0.854
Weak instrument	5.082**	4.495**	4.926**	5.676**	10.485**	7.221**	5.749**
Wu-Hausman test	1.419	0.003	5.169*	4.397*	5.323*	3.475'	0.95
Sargan test	15.473	10.918	10.604	5.455	13.899*	8.568	8.423
Moran's I resid.	-2.728	-1.89	-2.698	-2.976	-3.347	-2.572	-1.709

Note: *p<0.1; **p<0.05; ***p<0.01

4 Additional IV Regression Results

Table S2 replicates the models in Table S1, with proportion of pre-WWII Russian speakers (according to the 1931 Polish census) on the right-hand side. Despite larger standard errors due to the reduced sample size, the effects of violence are similar in direction and magnitude to those reported in the main text.

Table S3 reports the second and first stage results of the instrumental variable regression of 'pro-Russian' vote margin on red partisan control during World War II, with forest cover as an instrumental variable for partisan control. The standardized coefficients from these models appear in Figure 4 in the main text.

Table S4 replicates the models in Table S1, with OUN-UPA operations as the instrumented explanatory variable, and railroad access as the instrument. As the table reports, there is little evidence that railroads drove variation in rebel attacks, or that past rebel attacks drive voting today. The models do not pass the weak instrument test, and the second stage coefficients are insignificant.

Table S5 replicates the same models, with other Soviet forms of government violence (other than deportation) as the instrumented explanatory variable, and railroad access as the instrument. As the table reports, deportation had a far stronger impact on contemporary voting than these more general types of government operations – which do not have a discernible effect on the 'pro-Russian' vote in any election. There is also little evidence that railroads drove variation in these operations to the same extent. The difference in means for Soviet attacks located at below-average vs. above average distances from the railroad: 48 versus 44 operations, respectively, with a Kolmogorov-Smirnov p-value of .44. By contrast, the difference is 28 versus 21 for OUN-UPA attacks (p=.192) and 550 versus 448 for deportations (p=.066).

Table S2: **Instrumental variable regressions**, with 1931 ethno-linguistic composition as a covariate.

	Dependent variable: 'pro-Russian' vote margin						
	2014 Parl.	2014 Pres.	2012 Parl.	2010 Pres.	2007 Parl.	2006 Parl.	2004 Pres.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Second stage							
Soviet deportations	−0.075	−0.146**	−0.249***	−0.159*	−0.243***	−0.135	−0.079
	(0.067)	(0.069)	(0.082)	(0.096)	(0.083)	(0.096)	(0.088)
Russian speakers (1931)	45.611***	53.075***	4.902	42.647***	33.365***	31.477***	57.543***
	(9.334)	(10.317)	(12.061)	(12.772)	(10.352)	(12.031)	(10.794)
Covariates	Y	Y	Y	Y	Y	Y	Y
Oblast FE	Y	Y	Y	Y	Y	Y	Y
Moran eigenvectors	Y	Y	Y	Y	Y	Y	Y
Observations	201	202	194	201	199	200	202
R^2	0.909	0.888	0.855	0.820	0.885	0.843	0.872
Adjusted R^2	0.890	0.866	0.823	0.795	0.866	0.819	0.854
Weak instrument test	7.678**	7.339**	7.338**	5.566**	5.239**	7.072**	5.116**
Wu-Hausman test	0.577	2.662	4.624*	0.179	5.243*	0.729	0.007
Sargan test	11.251'	15.109'	10.661	7.987	8.116	6.688	7.884

Note: Intercept and control variables not shown. *p<0.1; **p<0.05; ***p<0.01

Table S3: **Instrumental variable regression results: partisan control**. Quantities are standardized coefficients, with standard errors in parentheses.

	Dependent variable: 'pro-Russian' vote margin						
	2014 Parl.	2014 Pres.	2012 Parl.	2010 Pres.	2007 Parl.	2006 Parl.	2004 Pres.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Second stage							
Partisan control	0.110***	0.147***	0.089**	0.088**	0.058	0.056	0.070*
	(0.039)	(0.045)	(0.045)	(0.042)	(0.035)	(0.041)	(0.039)
First stage							
Forest cover	0.051***	0.050***	0.051***	0.049***	0.051***	0.057***	0.055***
	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.018)	(0.017)
Covariates	Y	Y	Y	Y	Y	Y	Y
Oblast FE	Y	Y	Y	Y	Y	Y	Y
Moran eigenvectors	Y	Y	Y	Y	Y	Y	Y
Observations	217	217	207	217	215	216	218
R^2	0.862	0.829	0.826	0.854	0.897	0.870	0.871
Adjusted R^2	0.837	0.806	0.800	0.832	0.879	0.846	0.856
Weak instrument	9.037**	9.087**	8.101**	7.364**	8.442**	9.268**	6.724**
Wu-Hausman test	3.343'	4.187*	0.623	1.715	0.195	0.331	1.366
Sargan test	28.573'	23.656'	27.976	21.636	33.535*	19.698	38.29*
Moran's I resid.	-2.773	-2.404	-2.16	-2.202	-2.531	-2.67	-2.002

Note: Intercept and control variables not shown. *p<0.1; **p<0.05; ***p<0.01

Table S4: **Instrumental variable regression results: OUN-UPA violence**. Quantities are standardized coefficients, with standard errors in parentheses.

	Dependent variable: 'pro-Russian' vote margin						
	2014 Parl.	2014 Pres.	2012 Parl.	2010 Pres.	2007 Parl.	2006 Parl.	2004 Pres.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
OUN-UPA violence	0.168	−0.020	−0.189	−0.127	−0.053	0.019	−0.167
	(0.245)	(0.232)	(0.284)	(0.236)	(0.168)	(0.212)	(0.239)
Covariates	Y	Y	Y	Y	Y	Y	Y
Oblast FE	Y	Y	Y	Y	Y	Y	Y
Moran eigenvectors	Y	Y	Y	Y	Y	Y	Y
Observations	217	217	207	217	215	216	218
R^2	0.838	0.840	0.816	0.846	0.898	0.861	0.855
Adjusted R^2	0.813	0.817	0.789	0.825	0.881	0.842	0.837
Weak instrument test	0.228	0.289	0.316	0.434	0.14	0.219	0.247
Wu-Hausman test	0.657	0.003	0.213	0.22	0.02	0.008	0.462
Sagan test	20.472	34.204**	21.004*	16.63*	50.636*	38.766**	34.812**
Moran's I resid.	0.195	-2.251	-2.485	-3.071	-2.707	-1.682	-2.981
Note:	Intercept and control variables not shown. *p<0.1; **p<0.05; ***p<0.01						

Table S5: **Instrumental variable regression results: other Soviet violence**. Quantities are standardized coefficients, with standard errors in parentheses.

	Dependent variable: 'pro-Russian' vote margin						
	2014 Parl.	2014 Pres.	2012 Parl.	2010 Pres.	2007 Parl.	2006 Parl.	2004 Pres.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Other Soviet violence	0.025	0.022	0.077	0.011	0.009	0.057	0.085
	(0.050)	(0.063)	(0.065)	(0.058)	(0.047)	(0.055)	(0.057)
Covariates	Y	Y	Y	Y	Y	Y	Y
Oblast FE	Y	Y	Y	Y	Y	Y	Y
Moran eigenvectors	Y	Y	Y	Y	Y	Y	Y
Observations	217	217	207	217	215	216	218
R^2	0.868	0.838	0.832	0.853	0.899	0.870	0.872
Adjusted R^2	0.844	0.819	0.806	0.834	0.883	0.848	0.857
Weak instrument test	5.756**	6.533**	7.699**	7.227**	6.138**	8.027**	6.847**
Wu-Hausman test	0.061	0.009	0.288	0.054	0.373	0.013	1.53
Sargan test	36.143'	15.076	9.962	13.579	29.33*	15.204	16.55
Moran's I resid	-2.476	-1.803	-2.245	-1.996	-2.299	-2.318	-1.472
Note:	Intercept and control variables not shown. *p<0.1; **p<0.05; ***p<0.01						

5 Fuzzy RD: Assumptions and Robustness

5.1 Spatial "Discontinuities" in Deportation Levels

One of the assumptions that we invoke in the fuzzy regression discontinuity design (FRDD) is that the Soviet repression varied from district to district in part due to idiosyncratic variation in the vigilance and capability of the local NKVD and Communist party officials who were in charge of identifying whom to repress and how to execute the deportation operations. While in the text (section "Soviet Violence and Western Ukraine") we provide a lot of historical evidence to that effect, here we provide some numerical evidence to that effect as well. An observable implication of this assumption is that deportation levels should vary substantially across *spatially contiguous* districts (which we can plausibly expect to be similar in other characteristics). To put it in reverse, if the assumption we make does not hold in reality, then we should expect spatially contiguous districts to have roughly similar levels of deportation.

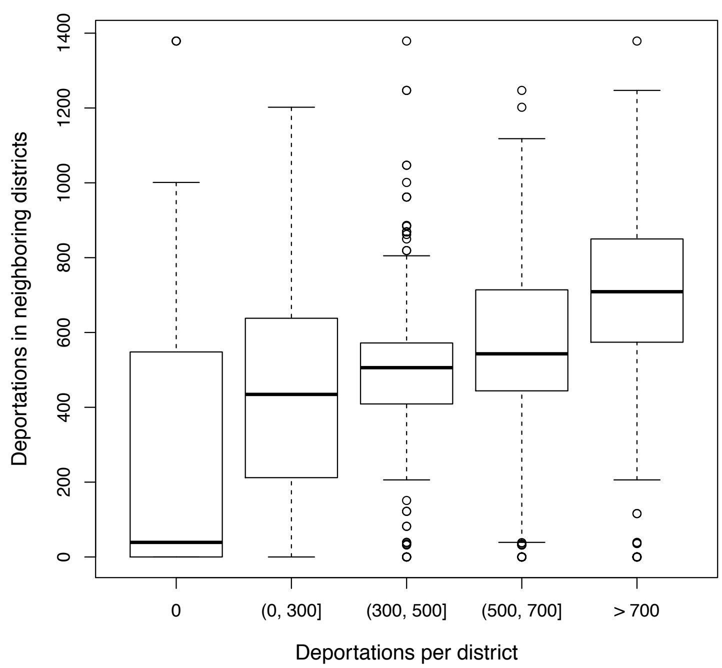


Figure S2: The variation of deportations across spatially proximate districts.

Figure S2 indicates that there is no strong evidence contradicting our assumption in the data. It shows how deportation levels across spatially contiguous districts vary for a given level of deportations. For example, if we take all districts with zero deportations (the left-most box in the box plot), we see that the median level of deportations in the districts that are spatially contiguous to them was highly variable, with the median of roughly 100, the lower quartile equal to 0, and the upper quartile above 200 deportations. Similarly, for each interval on the x-axist, the variation in the deportations in the neighboring districts is very large. In sum, even districts that are spatially proximate and due to that proximity are very likely to share many other characteristics, had highly contrasting levels of repression. Whether these contrasts are driven entirely by the differences in the administrative characteristics of the districts can be debatable, but in either case these patterns seem to support the case we are making for the fuzzy regression discontinuity design.

5.2 Balance Tests for the RDD Analysis

One concern for the fuzzy RDD analysis is that the rayon borders by the Soviets were drawn based on some pre-existing demographic and political characteristics that could also be driving the Soviet repression levels in the post-war days. While we have not encountered historical evidence supporting this concern, we believe it is still important to address it. To eliminate this concern is a formidable challenge, because it requires *settlement* level demographic and political data preceding the creation of the Soviet rayons in 1939 and 1940. The Polish census for 1931, from which we use some covariate information in our analyses, provides demographic measures only at the level of Polish administrative district (powiat). However, the Polish census of 1921 does provide some demographic information at the level of settlement, which we can use to check whether the demographic characteristics vary discontinuously across the rayon borders with contrasting levels of post-war deportations.

Settlements were geo-referenced using GeoNames database of settlement names. About one quarter of settlements could not be reliably matched and hence they are not present in the below analyses. For the balance tests, we select exactly the same set of rayons as in our main RDD analysis: neighboring rayons with *contrasting* levels of repression. We use the approach to RDD balance testing recently proposed in (de la Cuesta and Imai 2016), who argue that to asses whether the RDD design is well-balanced one needs to test for discontinuities at the cut-off of the forcing variable.

We consider six demographic covariates: three variables measuring the proportion of the three major religious affiliations (Orthodox, Greek catholic, and Roman catholic) per settlement and three variables measuring the proportion of ethnic groups (Polish, Jewish, and Ruthenians). As in our main analysis, the forcing variable is the distance from less repressive to more repressive rayon and we only include settlements located with 10 km of the rayon borders which enter the analysis (consistently with the main RDD analyses in the paper). To estimate the RDD effects and to plot them, we use the robust nonparametric procedure by(Calonico, Cattaneo and Titiunik 2014).

The results are show in Figure S3. The figures display the estimated RDD curves using quartic polynomials, in which we do not see clear indications of discontinuities. Inside each plot, we also report bias-corrected robust RD estimates of the discontinuity effect using local linear regression. Based on these coefficients and their p-values, we do not observe strong evidence that the covariates change discontinuously across the borders of rayons with contrasting levels of repression.

We should note that these balance tests also carry certain caveats: First, the 1921 census is biased in that it over-counted the Polish population (Kopstein and Wittenberg 2011), which might result in some bias, but we think it is somewhat unlikely because the rayon

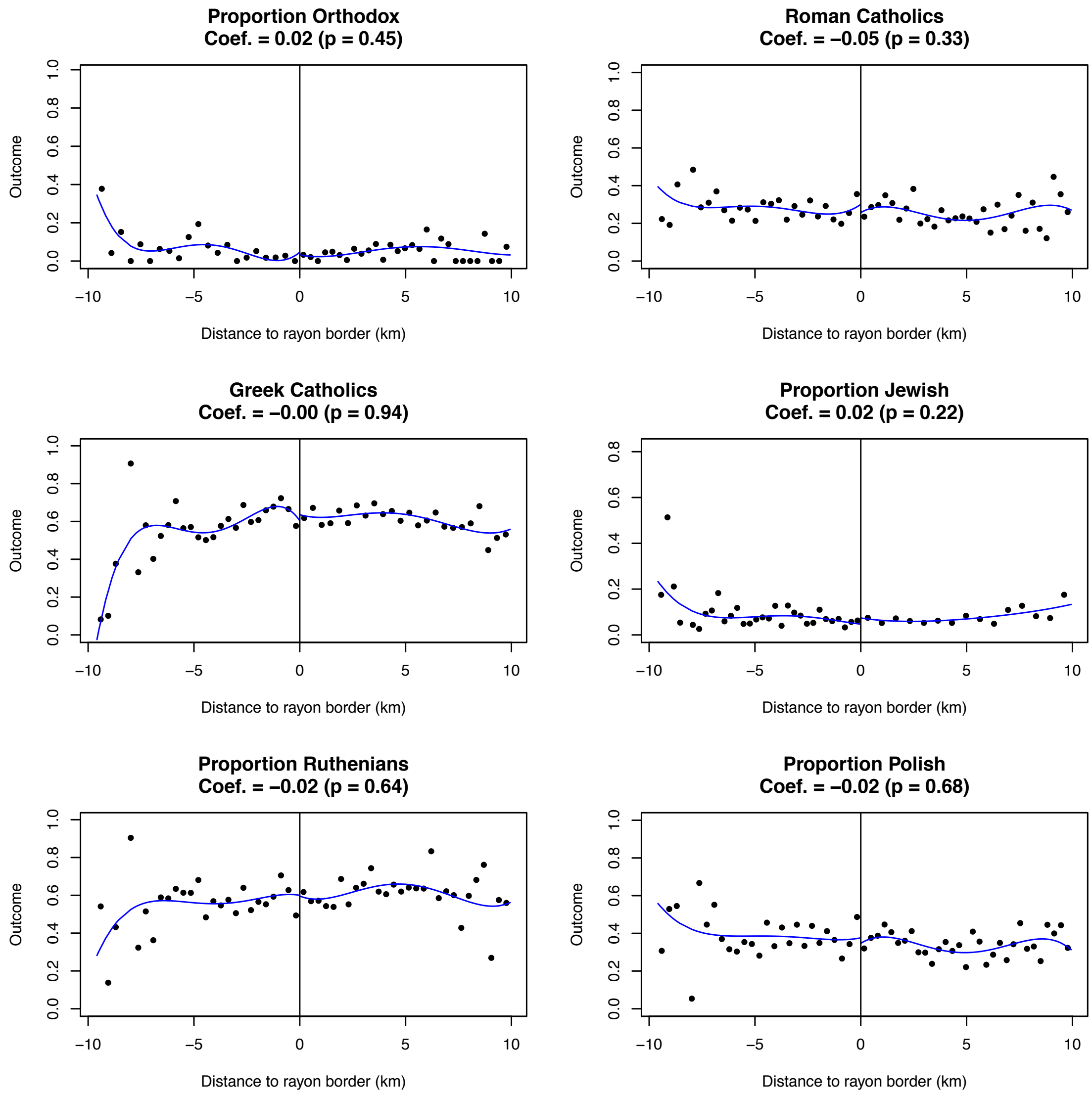


Figure S3: No evidence of discontinuities at rayon borders for the 1921 census data.

borders that we are using were drawn many years *after* the 1921 census, and it is not likely that the over-counting varied with respect to these future borders. Second, these RDD balance tests are potentially contaminated by the measurement error in geocoding the settlements, as their historical names have often changed from Polish into Ukrainian and our geo-coding method might have produced mistakes. Third, just because we have balance on these demographic covariates, does not imply that we would also have balance on other, unmeasured covariates.

Finally, we should also note that the plot indicate not only local but also global balance as the regression curves remain relatively flat even away from the cut-off point, which yields additional evidence in support for our rayon-level IV results.

5.3 Robustness to the Choice of the Contrast Cut-Off

The fuzzy RD analysis includes settlements nearby the border between two districts if those districts had sufficiently contrasting levels of historical repression (otherwise the comparison is not very meaningful). In the paper we use the rule that the two neighboring districts have contrasting levels of repression if one is one standard deviation above the sample mean and another is one standard deviation below sample mean. Here, we show that our results are robust to deviations from this rule. Figure 5.3 shows the estimated RD estimates for different cut-off rules from 0.5 standard deviations above and below the mean to 1.5 standard deviations above and below the mean. As we can see, in large set of cut-off values around 1, the results are similar to those reported in the paper: the effects of deportations are negative and significant at 95 percent confidence level.

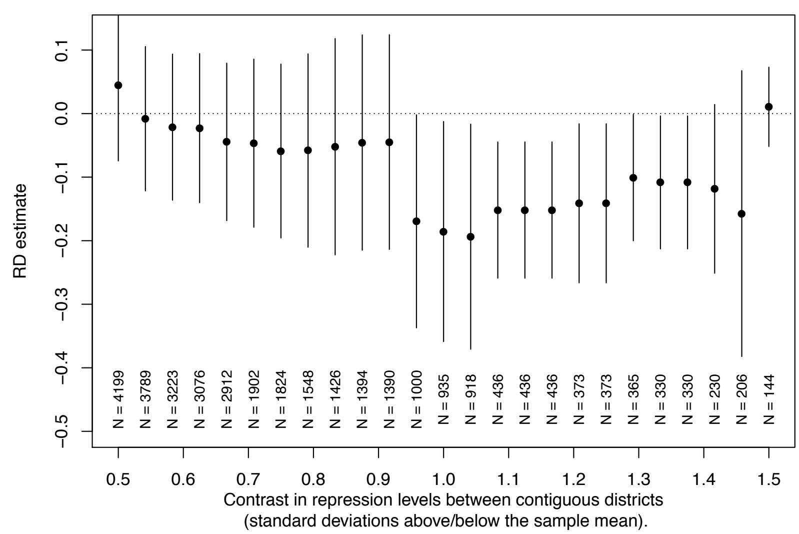


Figure S4: Fuzzy RD estimates and 95 percent confidence intervals for different cut-off rule used to select contiguous districts.

6 Placebo Tests for IV Regressions

A major concern for any instrumental variables analysis is the validity of the exclusion restriction, which requires that the instrument affects the outcome only through the treatment (Angrist and Pischke 2008). Since we use two instruments for two treatments in separate IV analyses, in our case the exclusion restriction requires that (1) distance to railways affects current voting patterns only through its effects on deportations and (2) forestation affects current voting patterns only through its effects on the red partisan control. The exclusion restriction can never be tested directly, but we can test some empirical implications that must be true in case the exclusion restriction is violated. Here we present a placebo test for our *Distance to railways* instrument.

The idea behind this placebo test is as follows: one region of our study, Zakarpattia oblast, was annexed to the Ukrainian Soviet Socialist Republic only in January of 1946. Consequently, it did not experience the first large wave of post-war deportations of 1944 and 1945, in contrast to other regions of western Ukraine. The map in Figure 1 of the paper shows that the Zakarpattia region (in the south western corner) had very few deportations compared to other regions. If the exclusion restriction were violated in our data, then distance to railways would affect current voting patterns through some mechanism other than deportations. The Zakarpattia region could therefore be treated as a placebo case: since deportations were extremely rare, the distance to railways could only have a negligible effect on the magnitude of deportations, which implies that a reduced form effect of distance to railways on current voting would indicate a violation of the exclusion restriction. More specifically, if the distance to railways has a *positive* association with current pro-Russian support, then the negative IV coefficients of deportations reported in the paper were negatively biased (thus, we would have *overestimated* the negative impact of deportations on pro-Russian support).

We conduct this placebo test using *polling station-level* voting data. We use a simple linear regression with polling station-level pro-Russian vote margin as a dependent variable. The independent variable is the shortest-path distance between the contemporary polling station and the post-war railways. As in the paper, we estimate the regressions for each election separately.

The results are reported in Table S6. The first panel of the table reports the results for "placebo" polling stations located in Zakarpattia region. For comparison, the right panel of the table reports coefficients of the same estimations but using polling stations that are located outside Zakarpattia, where the variation in deportations was significant. In the latter, non-placebo set of precincts, we expect the coefficient for *Distance to railways* to be positive because this would be consistent with our rayon-level IV results: the instrument (*Distance to railways*) has a negative effect on treatment (*Deportation*), and the treatment (*Deportations*) has a negative effect on the outcome (*Pro-Russian margin*), which implies that the reduced form effect of *Distance to railways*has a positive effect on *Pro-Russian margin*. As for the placebo precincts, if the exclusion restriction is violated in a way that negatively biases our IV results, then the coefficient for *Distance to railways* would be positive as well. A zero or negative coefficient would indicate that either the exclusion restriction is not likely to be

	Placebo precincts (Zakarpattia)			Non-placebo precincts (other regions)		
	Coef.	S.E.	p-value	Coef.	S.E.	p-value
2004 Presidential	−0.39	0.19	0.04	0.21	0.02	0.00
2006 Parliamentary	−0.26	0.14	0.06	0.23	0.02	0.00
2007 Parliamentary	−0.11	0.13	0.39	0.30	0.02	0.00
2010 Presidential	−0.36	0.21	0.09	0.32	0.03	0.00
2012 Parliamentary	0.14	0.12	0.26	0.51	0.02	0.00
2014 Parliamentary	−0.07	0.07	0.32	0.16	0.01	0.00
2014 Presidential	−0.04	0.06	0.54	0.15	0.01	0.00

Table S6: Reduced form precinct-level regressions (OLS coefficients). The dependend variable is pro-Russian vote-margin and the independent variable is the shortest-path distance between the current polling station and post-WWII railways (in km).

violated or that it is violated in a way that attenuates the negative IV estimate of the causal effect of deportations on pro-Russian support.

As we can see from the results in Table S6, for the placebo precincts, the reduced from effect is not significant at 95 the percent level in all but 2004 elections, and in 2004 elections it is actually negative. This indicates that *Distance to railways* is certainly not positively correlated with current pro-Russian support, which leads us to conclude that we do not see evidence of the exclusion restriction being violated in these data. Consistent with our rayon-level IV results, the reduced form effect *Distance to railways* on *Pro-Russian support* at the precinct level is positive and statistically significant in the set of non-placebo cases, lending further support for our IV results.

7 Caetano Exogeneity Test

Recently, Caetano (2015) has proposed a treatment exogeneity test, for cases where the treatment variable has bunching at certain values. In the case of our data, there is very clear bunching of deportation values at zero – about 10 percent of rayons did not experience deportations (Figure S5). This allows us to conduct the said exogeneity test for the *Deportations* variable (note we cannot do this for the binary *Partisan control* variable).

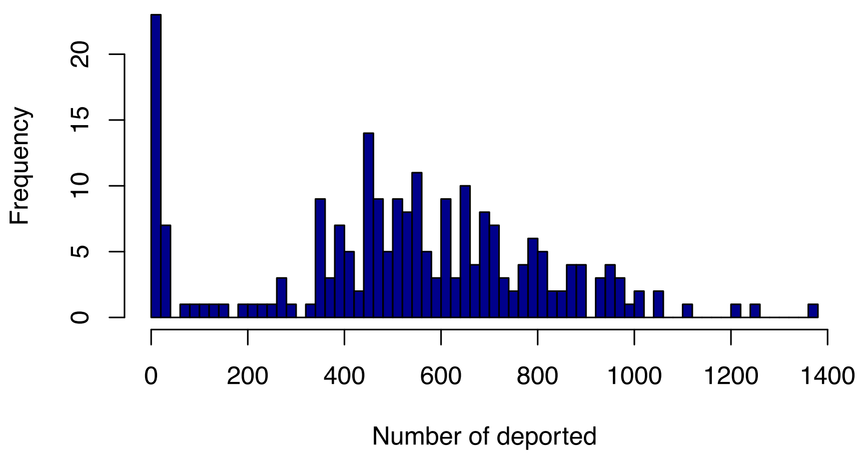


Figure S5: Distribution of deportations shows clear bunching at zero.

The key idea behind the Caetano (2015) exogeneity test is as follows. Suppose we assume (not unreasonably) that the effect of the treatment on the outcome varies continuously. In our case, this assumption implies that increasing historical deportation levels by a few persons cannot *discontinuously* increase the current support levels for pro-Russian parties. If the treatment variable is exogenous, therefore, we should expect no significant difference in pro-Russian support across rayons that experienced no deportations versus rayons that experienced very minor levels of deportations (the minimum *positive* number of deportees in our data is 32). If, however, there is discontinuous change in current pro-Russian support as we move from rayons with no deportations to rayons with, say, 50 or 100 deportations, then some unobserved factor may be driving this discontinuous change in the outcome variable.

The test proceeds by estimating the quantity θ, representing the discontinuous change in the expected value of the response variable at the limit of the explanatory variable x:

$$\theta = \lim_{x \to 0} \mathbb{E}[\mathbb{E}[Y|X = 0, X] - Y|X = x],$$

where X represents the explanatory variable of interest (in this case, *Deportations*), Y represents the outcome variable (pro-Russian electoral support, scaled in terms of standard deviations from the mean), and Z represents the set of control variables (this includes all the control variables used in our baseline IV estimations). To implement the method, one needs to pick the bandwidth for the values of the treatment x used to estimate the value of θ. Ideally, we would like to pick as small a bandwidth as possible. In reality, however, we are constrained by the small sample, which can lead to highly unstable estimates when the chosen bandwidth is small.

Bandwidth (# deported)	Estimate ($\hat{\theta}$)	St. Error
50	−27.94	3.52
75	−25.75	3.51
100	.417	.59
125	−.40	.47
150	−.59	.42

Table S7: Exogeneity test results for different bandwidths.

Table S7 reports estimated values for different bandwidths (measured in terms of the number of people deported), using the method and software by Caetano (2015). For small bandwidths, the estimated $\hat{\theta}$ is large and negative, indicating that if any confounders exist, they are likely to *positively* bias the estimated association between deportations and pro-Russian voting, because, according to the results at bandwidths of 50 and 75 deported people, pro-Russian support discontinuously *drops* at the minimal values of the treatment, meaning that the negative relationship between treatment and outcome is *underestimated*.

We should caution, however, that the $\hat{\theta}$ estimates are very large for small bandwidths perhaps due to a small number of observations in these bandwidths. However, the key point here is that the estimates of $\hat{\theta}$ are not positive and statistically significant either for small or larger bandwidths, as the standard errors for bandwidths of 100, 125, 150 deportees are either larger or very similar in magnitude to the estimate itself.

The simple standardized OLS estimate of the effect of deportations on pro-Russian support (using data pooled from all elections), after adjusting for controls and regional fixed effects, is -0.061 with the standard error of .012 ($p < 0.01$) (see Appendix 8 for details). This OLS effect is consistent in its direction to the IV estimates reported in the paper, but it is substantially smaller than the IV estimates. The results in Table S7 indicate why that could be the case: there is some confounding in the data, even after adjusting for covariates, that *positively* biases the OLS coefficient. In conclusion, even without the IV or RDD analysis (which require their own assumptions), a simple OLS regression with covariate adjustment also yields the negative and statistically significant effect of deportations on current pro-Russian support, and as the results of this section indicate, this is likely to be an underestimate, which is corrected in the IV and RDD analyses.

8 Population Size as Alternative Mechanism

One alternative explanation of our results is that they are driven by the rayon-level population size and/or urbanization. The mechanism would be as follows: more populated places saw greater density of railways in the pre-WWII times, which exposed them to greater levels of deportations, but down the line more densely populated places have also become less pro-Russian (we thank one of the reviewers for bringing up this point). In this way, our IV analysis would violate the independence assumption (Angrist and Pischke 2008).

The problem is best addressed by separating it in two parts – the impact of historical population size and the impact of current population size. In Table S1 of Appendix 3, we show that our IV results hold even when we adjust for pre-war urbanization level (a proxy for population size). We cannot address the second concern in the same fashion by controlling for the current population size in the IV regressions because it would mean that, in the first stage regressions, we would be using *future* urbanization to predict WWII deportation levels, leading to bias due to post-treatment adjustment (Rosenbaum 1984). However, we can evaluate the plausibility of this alternative mechanism by investigating the reduce-form relationship between pre-WWII deportations and current pro-Russian voting with and without using post-treatment adjustment for current population size.

We estimate the following fixed effects regression:

$$Margin_{i,t} = \beta_0 + \beta_1 Deportations_i + \beta_2 Log(Population_{i,t}) + \eta_{j[i]} + u_t + \epsilon_{i,t},$$

where $Margin_{i,t}$ is pro-Russian margin in rayon i at election t, $Deportations_i$ is deportation level at rayon i, $Population_{i,t}$ is population in rayon i at election t (which we measure by the number of registered voters in the rayon), $\eta_{j[i]}$ and u_t is the oblast-level and election-level fixed effect, respectively. The idea here is that if the purported mechanism is driving the results, then two things should happen: the effect of deportations should be smaller after we control for current population and the effect of current population on pro-Russian vote should be negative.

Table S8 reports the results with the coefficients normalized to represent the effects in terms of standard deviations from the mean. Without controlling for present-day population, the reduced-form effect of *Deportations* is negative and highly significant, though substan-

	(1)	(2)
Deportations	-0.061^{***}	-0.081^{***}
	(0.012)	(0.015)
Log(population)		0.049^{***}
		(0.018)
Observations	1,507	1,506
R^2	0.869	0.871
Adjusted R^2	0.868	0.870
Residual Std. Error	0.363 (df = 1492)	0.361 (df = 1490)
F Statistic	707.635^{***} (df = 14; 1492)	670.600^{***} (df = 15; 1490)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table S8: Reduced-form regressions. Intercept, oblast-level and election-level fixed effects not reported. Standard errors in the parentheses are clustered at the election level.

tially smaller than the IV or the fuzzy RDD effect. Once we control for current population (approximated by the number of registered voters), the coefficient for deportations *increases* and, furthermore, the coefficient for *Log(population)* is *positive* and statistically significant. Both of these findings are *not* consistent with the purported proposition that deportation effects are driven by differential population levels across the rayons. Post-treatment adjustment for current population increases the estimated effect of deportations.

9 Test of the "Ethnic Composition" Mechanism

This section reports two sets of results. The first (Table S9) tests whether Soviet-era violence helps predict the contemporary linguistic composition of western Ukrainian districts, measured as the difference between the percentages of Russian and Ukrainian speakers, according to the 2001 census. The second (Table S10) is a replication of the models in Table 3 in the main text, with the difference between Russian and Ukrainian speakers in 2001 included as a covariate. Table S9 suggests that Soviet violence did not affect contemporary linguistic composition. Table S10 shows that the effect of violence remains after this post-treatment adjustment.

Table S9: Instrumental variable regressions, with ethno-linguistic composition in 2001 (percent Russian-speakers minus percent Ukrainian-speakers) as dependent variable.

	Dependent variable:	
	% Russian-speakers - % Ukrainian-speakers (2001)	
	(1)	(2)
Soviet deportations	−0.116	
	(0.102)	
Partisan control		−0.040
		(0.051)
OUN-UPA violence	0.003	0.002
	(0.002)	(0.002)
Regional FE	Y	Y
Moran eigenvectors	Y	Y
Observations	218	218
R^2	0.758	0.760
Adjusted R^2	0.703	0.709
Weak instrument test	8.442**	12.152**
Wu-Hausman test	0.594	0.599
Sagan test	21.546*	17.939
Moran's I resid.	-2.91	-2.731
Note:	Intercept and control variables not shown. *p<0.1; **p<0.05; ***p<0.01	

Table S10: IV regression results, with 2001 Russian language as additional covariate.

	Dependent variable: 'pro-Russian' vote margin						
	2014 Parl.	2014 Pres.	2012 Parl.	2010 Pres.	2007 Parl.	2006 Parl.	2004 Pres.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Second stage							
Soviet deportations	−0.131**	−0.231***	−0.258***	−0.060	−0.190***	−0.196***	−0.203***
	(0.059)	(0.088)	(0.086)	(0.084)	(0.056)	(0.066)	(0.076)
Ethno-linguistic composition (2001)	0.026***	0.023***	0.021***	0.036***	0.026***	0.026***	0.022***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Oblast FE	Y	Y	Y	Y	Y	Y	Y
Moran eigenvectors	Y	Y	Y	Y	Y	Y	Y
Observations	217	217	207	217	215	216	218
R^2	0.916	0.872	0.874	0.903	0.930	0.898	0.892
Adjusted R^2	0.899	0.855	0.853	0.891	0.916	0.880	0.878
Weak instruments	7.221**	6.441**	5.364**	5.844**	9.872**	7.516**	5.393**
Wu-Hausman test	1.136	3.753'	3.287'	0.54	2.798'	3.88'	2.732'
Sagan test	12.169	8.655	12.423	6.457	8.698	6.839	5.288
Moran's I resid.	-2.73	-1.955	-2.244	-1.687	-3.436	-2.856	-2.294
Note:	Intercept and control variables not shown. *p<0.1; **p<0.05; ***p<0.01						

References

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Caetano, Carolina. 2015. "A test of exogeneity without instrumental variables in models with bunching." *Econometrica* 83(4):1581–1600.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82(6):2295–2326.

de la Cuesta, Brandon and Kosuke Imai. 2016. "Misunderstandings About the Regression Discontinuity Design in the Study of Close Elections*." *Annual Review of Political Science* 19:375–396.

Kopstein, Jeffrey S and Jason Wittenberg. 2011. "Deadly communities: Local political milieus and the persecution of Jews in occupied Poland." *Comparative Political Studies* 44(3):259–283.

Rosenbaum, Paul R. 1984. "The consquences of adjustment for a concomitant variable that has been affected by the treatment." *Journal of the Royal Statistical Society. Series A (General)* pp. 656–666.