Online Appendix

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

Contents

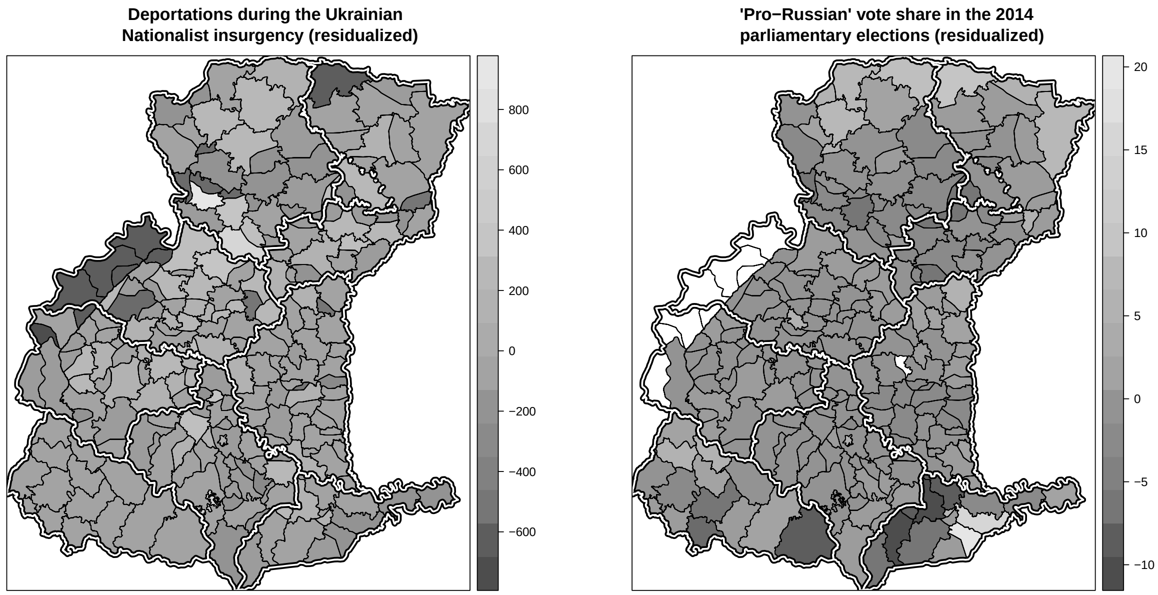
1 Classifying Political Parties and Candidates

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.

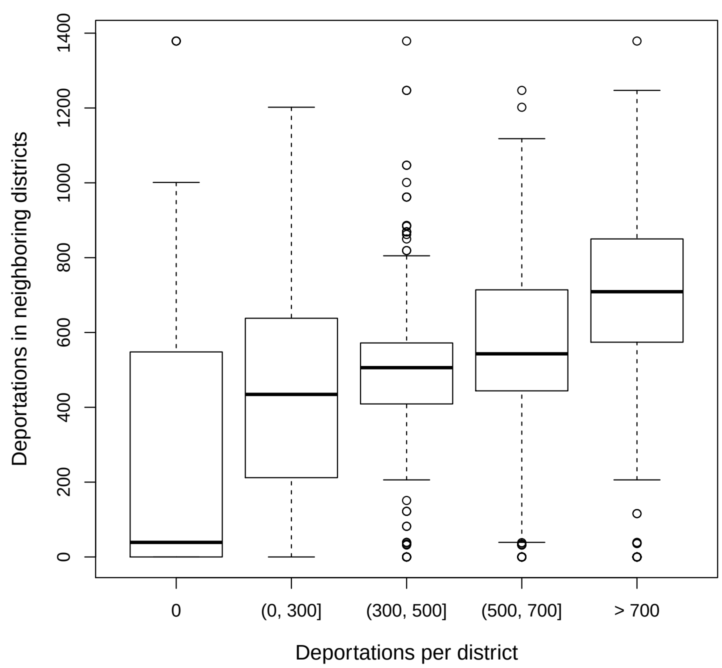


Figure S2: The variation of deportations across spatially proximate districts.

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

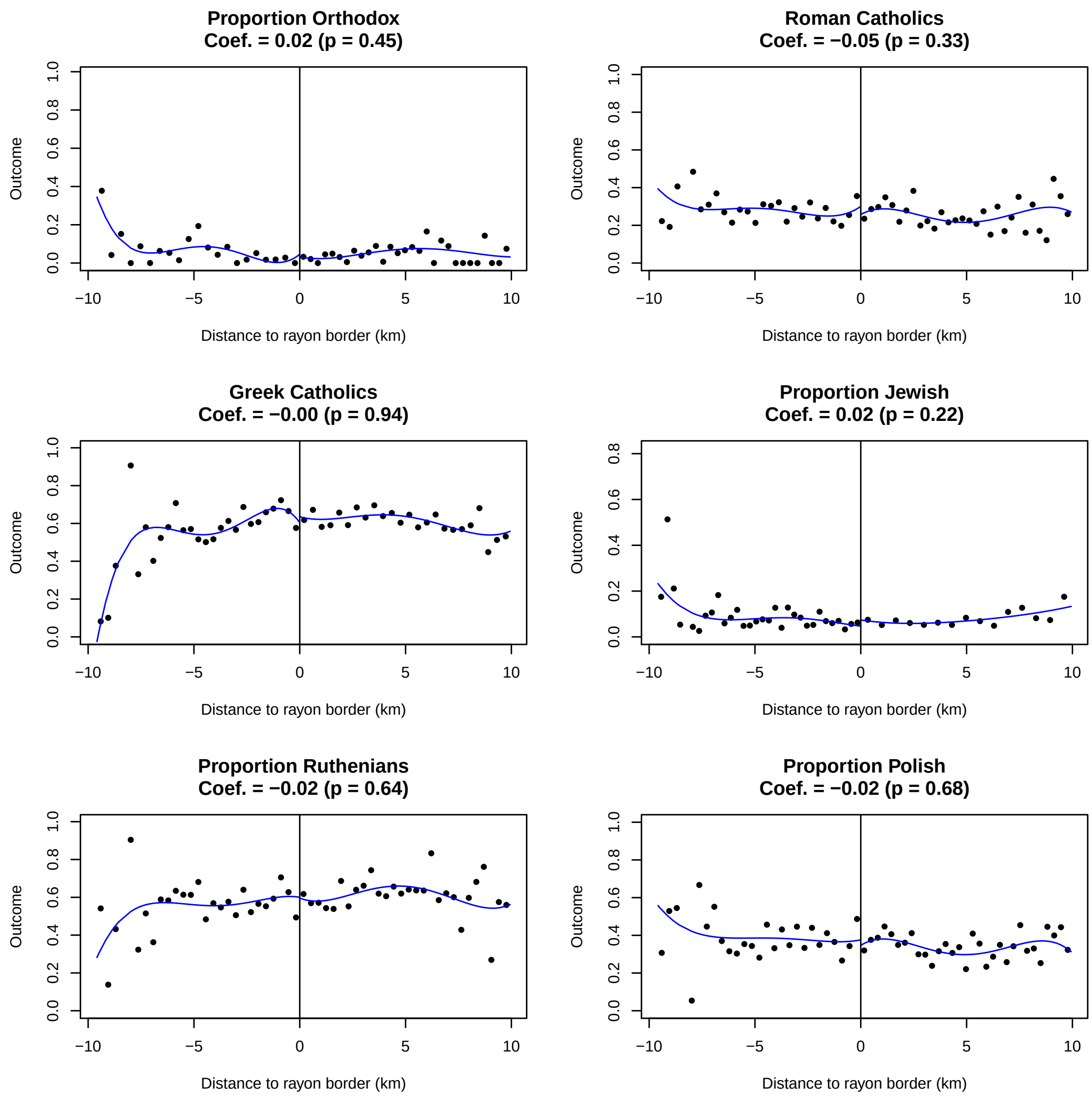


Figure S3: No evidence of discontinuities at rayon borders for the 1921 census data.

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.

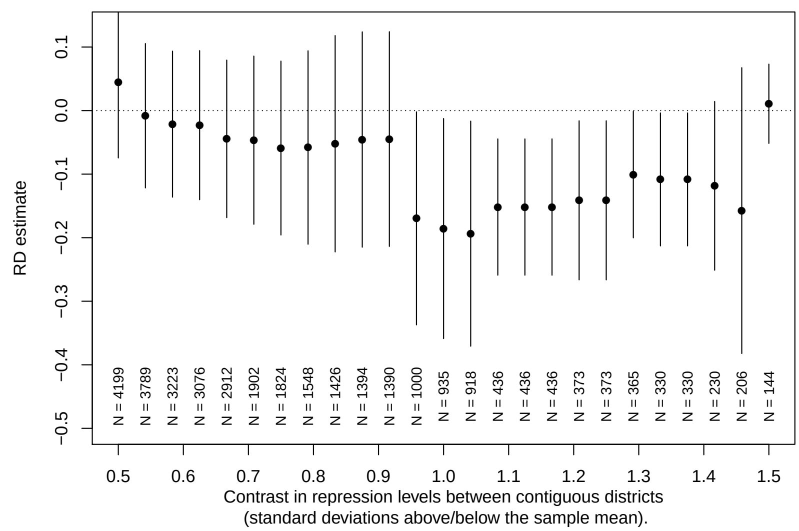


Figure S4: Fuzzy RD estimates and 95 percent confidence intervals for different cut-off rule used to select contiguous districts.

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).

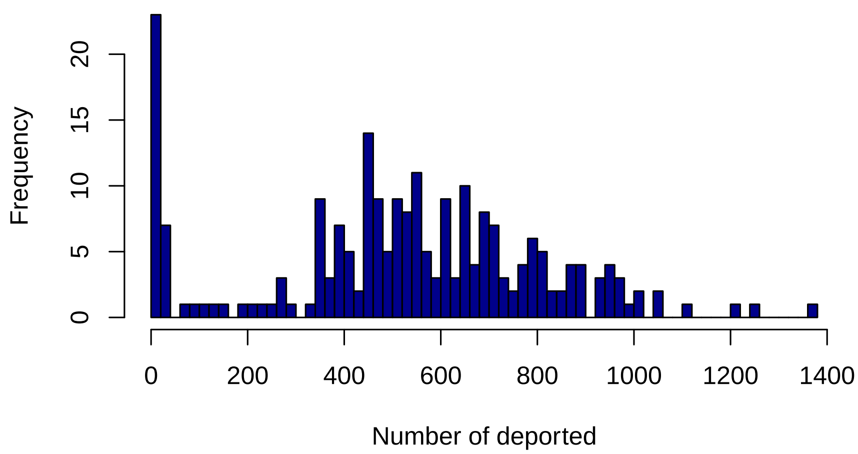


Figure S5: Distribution of deportations shows clear bunching at zero.

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.