Appendix

In this appendix, we provide additional details on how we used ArcGIS to create our global dataset of directed dyads, how we created measures of the proportion of border areas that are mountainous and the number of border crossings in each border area, and the results of additional models.

Creating the global dataset

In the main text, we stated that we included all contiguous directed dyads and the dyads with the four physically closest countries for each country, as calculated by ArcGIS. We obtained the list of contiguous directed dyads from the Correlates of War project. For non-contiguous directed dyads, we pursued a longer process. First, we counted the number of countries (n) that share a border with each country. We then generated neighbor lists using spatial weights matrices for the n+4 neighbors of each country. This allowed us to identify the four closest non-contiguous countries for each country. Our directed dyad list of contiguous and non-contiguous directed dyads goes from 1980-2016, so we accounted for state entry and exit in the international system. Again, we relied on the Correlates of War project for data on state entry and exit. Our analysis in this paper is confined to 1990-2016, so we only used directed dyads for this time period.

Creating measures of mountainous borders and border crossings

In order to account for geographic features of borders, we began by obtaining the locations of all land borders. To do this, we: 1) Used the Intersect function in ArcGIS on the global polygon shapefile of countries from the cshapes package to create a line shapefile of country borders; 2) Removed water borders and borders that did not exist before 2011; 3) Added Correlates of War country codes for each border and verified that we had all directed dyads represented as border lines.

Then, we turned all of our borders into "border areas." To accomplish this, we added a 5 km buffer around all border lines. This step in ArcGIS reduced the directed dyads down to border areas that only had one of the two directed dyads, so we had to keep this in mind at later steps.

Our *mountain* variable is a continuous measure from zero to one. We use a binary measure of whether 1 km grid cells are mountainous. This measure comes from the global raster data produced by Kapos, Rhind et al. (2000). Then, we calculated proportion of the border area that is mountainous. For example, a value of 0.4 means that 40% of the border area is mountainous.

Our *border crossing stations* variable is a count of the number of border crossings in a border area. Border crossing stations were identified by identifying locations where roads intersect borders.[1] This border crossing variable is thereby going to miss border crossings on roads that are not captured in existing spatial datasets. For a variety of reasons, road network shapefiles miss the existence of many roads (Center for International Earth Science Information

[1] Data was downloaded from: http://worldmap.harvard.edu/geoserver/wfs?outputFormat=SHAPE-ZIP&service=WFS&request=GetFeature&format_options=charset%3AUTF-8&typename=geonode%3Aborder_crossings_phv&version=1.0.0

Network - CIESIN - Columbia University and Information Technology Outreach Services - ITOS - University of Georgia 2013). We expect therefore that our count of border crossings will tend to undercount crossings, and it will best reflect the existence of long-term crossing points.

Estimating the effects of border geography on barrier existence and refugee flows

In Table A1 and Table A2 below, we estimate poisson models to examine the effects of border geography on barrier existence and refugee flows. For our models of barrier existence in Table A2, we estimated models of all three of our barrier variables.

These models show that *Mountainous Territory* decreases the likelihood of having a barrier and decreases the size of refugee flows, on average. This supports the view that mountains increase the difficulty of building barriers, and they deter people from crossing borders. We do see in our case study of the United States-Mexico border that it is possible that if countries with mountainous borders build barriers, then they will try to use the natural barrier of mountains as a complement to the human-made barrier. This logic would merit additional analysis though.

We also see that *Border Crossing Stations* increase the likelihood of having a barrier and increases the size of refugee flows, on average. This supports our view that borders with more official crossing points are easier to cross and are easier locations for countries to place barriers.

Then, we find that *Land Boundary Length* increases the likelihood of having a barrier and increases the size of refugee flows, on average. The only exception is that boundary length is not significantly related to refugee flows when using the Hassner and Wittenberg (2015) measure. This also supports the view that longer borders are easier to cross and are easier locations for countries to place barriers.

Table A1: Poisson regression for border geography on refugee flows

VARIABLES	(1) Logged Refugee Flows
Mountainous Territory	-0.167***
	(0.033)
Border Crossings	0.009***
	(0.001)
Logged Boundary Length	0.042***
	(0.013)
Lagged and Logged Refugee Flows	0.217***
	(0.003)
Civil War Intensity, Origin	0.301***
	(0.018)
Polyarchy, Origin	-0.653***
	(0.061)
Genocide, Origin	0.006
	(0.011)
Political Terror Scale, Origin	0.265***
	(0.013)
Logged GDP per Capita, Origin	-0.181***
	(0.010)
Logged Distance between Countries	-0.081***
	(0.014)
Common Language	0.184***
	(0.022)
Colonial Tie	0.453***
	(0.044)
Constant	-0.546
	(0.192)
Observations	18,100

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2: Logistic regression for border geography on barrier existence

VARIABLES	(1) Hassner & Wittenberg	(2) Poast and Carter	(3) Avdan and Gelpi
Mountainous Territory	-1.603***	-1.026***	-1.090***
	(0.136)	(0.143)	(0.124)
Border Crossings	0.017***	0.013***	0.009***
	(0.002)	(0.003)	(0.002)
Logged Land Boundary Length	0.062	0.146***	0.262***
	(0.044)	(0.051)	(0.045)
Lagged Civil War Intensity, Origin	0.177**	-0.014	0.191***
	(0.075)	(0.084)	(0.074)
Lagged Polyarchy, Origin	-2.540***	-1.827***	-2.486***
	(0.187)	(0.213)	(0.176)
Lagged Genocide, Origin	-0.679***	-0.727***	-0.472***
	(0.154)	(0.209)	(0.111)
Lagged Political Terror Scale, Origin	0.134***	0.192***	0.032
	(0.044)	(0.049)	(0.043)
Lagged and Logged GDP per Capita, Origin	0.129***	0.004	0.215***
	(0.032)	(0.037)	(0.031)
Logged Distance between Countries	-0.076	0.008	0.033
	(0.049)	(0.057)	(0.050)
Common Language	-0.316***	0.066	-0.150*
	(0.084)	(0.091)	(0.079)
Colonial Tie	-0.718***	0.552***	-0.044
	(0.227)	(0.173)	(0.168)
Constant	-3.434***	-5.131***	-7.279***
	(0.670)	(0.769)	(0.669)
Observations	18,172	18,172	18,172

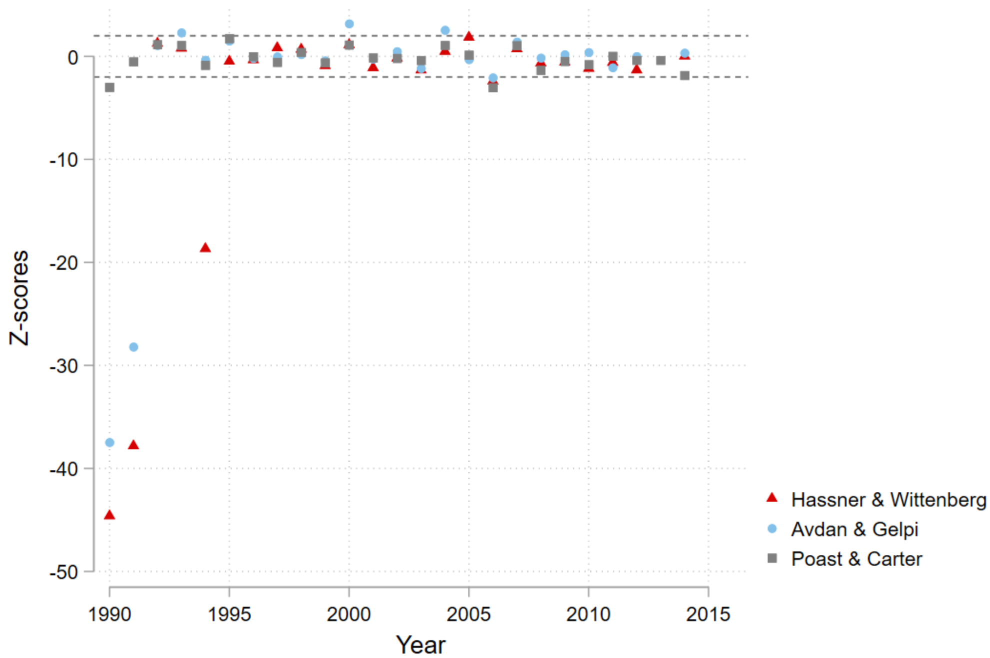
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Negative binomial results

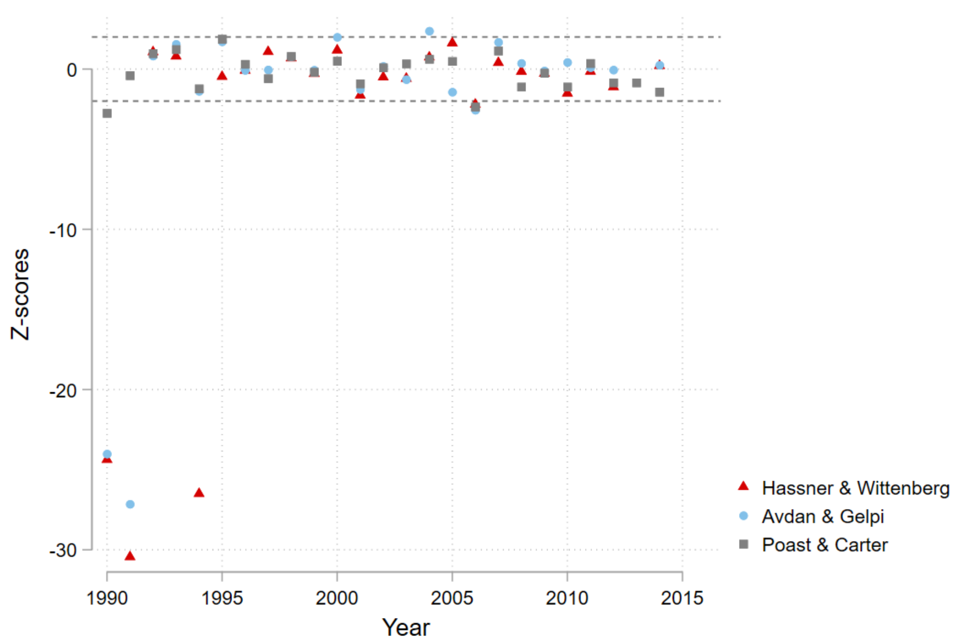
In Figure A1 below, we show the results of negative binomial regressions. This is the same model specification as Figure 8 in the main text, except that this is negative binomial instead of poisson. Here, we see that our negative binomial results are very similar to our poisson results. Aside from the early 1990s, there is not a statistically significant relationship between barriers and refugee flows.

Figure A1: Negative binomial results for all dyads and contiguous dyads

All dyads



Contiguous dyads



OLS, Tobit, and OLS-IV results

In Figures A2, A3, and A4 below, we display model results from OLS, Tobit, and OLS with instrumental variables. These linear models have some differences with our event count models. We do not discuss them in the main text because a normal distribution does not fit the distribution of our dependent variable as well as a poisson distribution.

Here, the OLS models show that most model specifications yield an insignificant relationship between barriers and refugee flows. When this relationship is statistically significant, it is negative. This correlation, however, appears to be the result of endogeneity, as we can see in Figure A3. Figure A3 shows that with the inclusion of an instrumental variable, *Logged Land Boundary Length*, there are no model specifications with a statistically significant relationship between barriers and refugee flows. In an effort to align the assumed distribution of the dependent variable with its actual distribution, we also estimated tobit models in order to create Figure A4. Generally, these models mostly fail to detect a statistically significant relationship between barriers and refugee flows.

Figure A2: OLS results

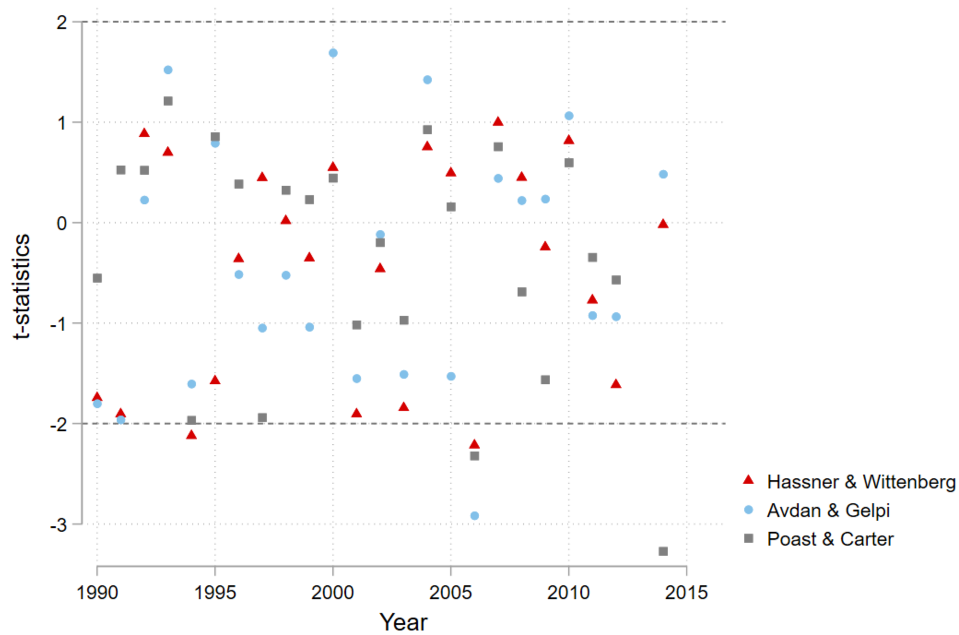


Figure A3: OLS with instrumental variable

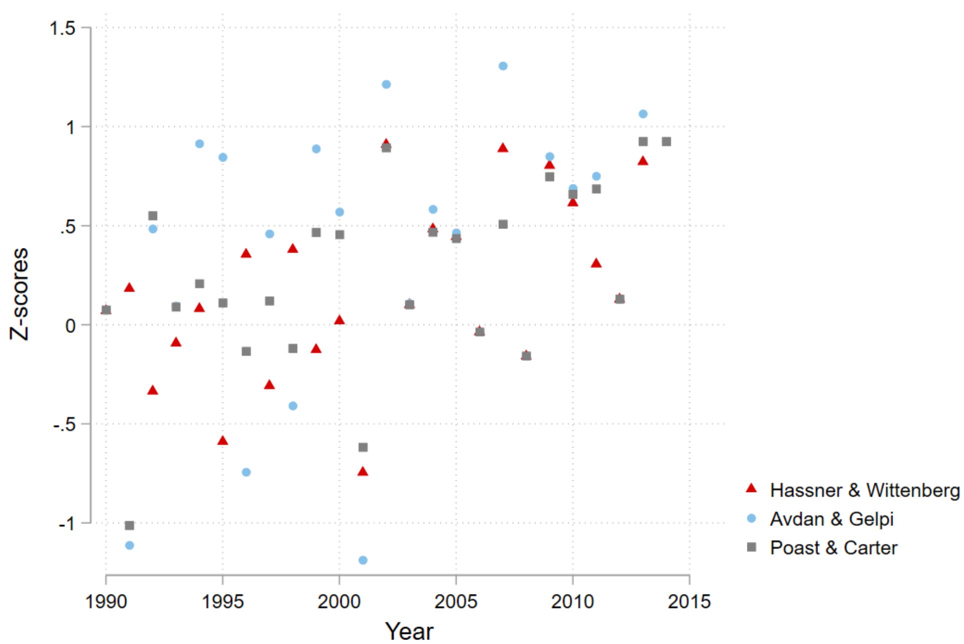


Figure A4: Tobit results