Supplementary Online Appendix

Volha Charnysh

Assistant Professor of Political Science

Massachusetts Institute of Technology

charnysh@mit.edu

A Background on post-WWII population transfers

A.1 Description of the main population groups in the resettled territories.

Indigenous/autochthonous population (*autochthoni*):

At the end of the war, about 1.2 million people living in German territories declared Polish nationality. Approximately one million (83%) of them passed verification procedures in the late 1940s and remained in Poland. The verification committee tasked with ascertaining "Polishness" usually consisted of a local governor or foreman (as chairman); two representatives of the People's Council; some representatives of the autochthonous population who had already proven their allegiance to Poland; representatives of Polish Western Union (*Polski Związek Zachodni*), an anti-German patriotic organization in the western regions; a Catholic priest; a teacher; a head of the District Office of Public Safety (*Powiatowy Urząd Bezpieczeństwa Publicznego*); and a head of the district police station (*Milicja Obywatelska*) (Lach 1978). The procedure was far from objective. As Dariusz Stola (2010, 67) notes, considerable discretion exercised by local officials and the circumstances in which the local population chose to undergo verification resulted in "hundreds of thousands of cases in which pro-Polish autochthons were expelled, while many pro-German autochthons were verified as 'Polish' and kept in place, sometimes against their will."

The autochthonous population was employed predominantly in agriculture and typically spoke little Polish. In fact, the majority were neither Polish nor German, and identified with smaller ethnic groups: Catholic Kashubians in the northwest; Protestant Mazurians and Catholic Warmiaks in the north; and Catholic Silesians in the southwest. Most hoped to stay in their homes upon hearing about the fate of the expelled Germans, who suffered from hunger and the lack of housing. Later decades would see a mass exodus of the indigenous population, who sought to return to Germany having experienced life under communist rule.

Repatriates from the USSR (*repatrianci*):

The eastern territory annexed from Poland by the Soviet Union (*Kresy*) was divided between the newly created Lithuanian, Belarusian, and Ukrainian Soviet Socialist Republics (SSRs). According to the 1931 Polish census, 10.7 million people lived in this area before the war, and 36% of them identified as Polish. Although the pre-war statistics exaggerated the size of the Polish population and did not account for the demographic changes caused by the Holocaust and Soviet deportations, the annexation meant that millions had to be resettled. Jan Czerniakiewicz (1987, 30) estimates that in 1944, Poles and Jews eligible for repatriation in line with the Yalta and Potsdam agreements numbered about 2.2-2.7 million.

The population from western Ukraine was the first to relocate as they fled the attacks of the Ukrainian Insurgent Army (UPA): statistics indicate that 789,982 people (92.2% of those eligible for repatriation) were resettled, accounting for 51.88% of all USSR repatriates. In Belarus and Lithuania, by contrast, interethnic boundaries were more ambiguous and intergroup relations less conflictual, which reduced incentives to emigrate. In addition, the local Soviet authorities sought to prevent the Catholic population in rural areas from leaving during the sowing season. As a result, 53% and 52% of eligible population emigrated from Belarus and Lithuania, respectively.

The USSR repatriates came from predominantly rural areas; some 60.81% lived in rural areas prior to relocation. This migrant group was representative of the general population in Kresy: 33.38% of them were peasants; 16.64% were workers; 11.63% craftsmen; and just 23% were white-collar employees. Most transports of USSR "repatriates" were directed to the territories obtained from Germany. A total of 1.9 million were settled there by 1948. Only a small group was settled in the old Polish lands, numbering 250,000 in rural areas and 300,000 in urban areas (Czerniakiewicz 1987).

Settlers from Central Poland (*osadnicy*):

Another 2.2 million migrants came from central Poland. These migrants left homes voluntarily and are thus not entirely representative of the general population. They came predominantly from rural areas, but also included the inhabitants of small towns in central Poland and of large cities destroyed by war, including Warsaw, Poznan, Bialystok, and Grudziadz. Overall, approximately 51.2% of voluntary migrants came from what used to be the Russian partition, which included Warsaw and suffered the most destruction during WWII. Another 25.2% of migrants came from the Austrian partition and the remaining 23.6% originated in the Prussian partition.

Some groups had an easier time relocating than others. One included people who had lost everything during the war. Not only did they have nothing to lose by moving west, but they also received preferential treatment by the authorities. Under Circular No. 22 issued in March 1946, farmers from regions damaged during the war were given priority for resettlement. Another group included inhabitants of the areas bordering the German territories, as well as areas assigned to act as patrons for specific localities in the west. For example, a coal mine in prewar Polish Upper Silesia was tasked with helping to revive its Lower Silesian counterpart by sending volunteer crews (Blusiewicz 2015).

These migrants also had little agency in choosing their destinations in western Poland, though they sometimes left when dissatisfied with their assignments of farms or the living conditions on the frontier.

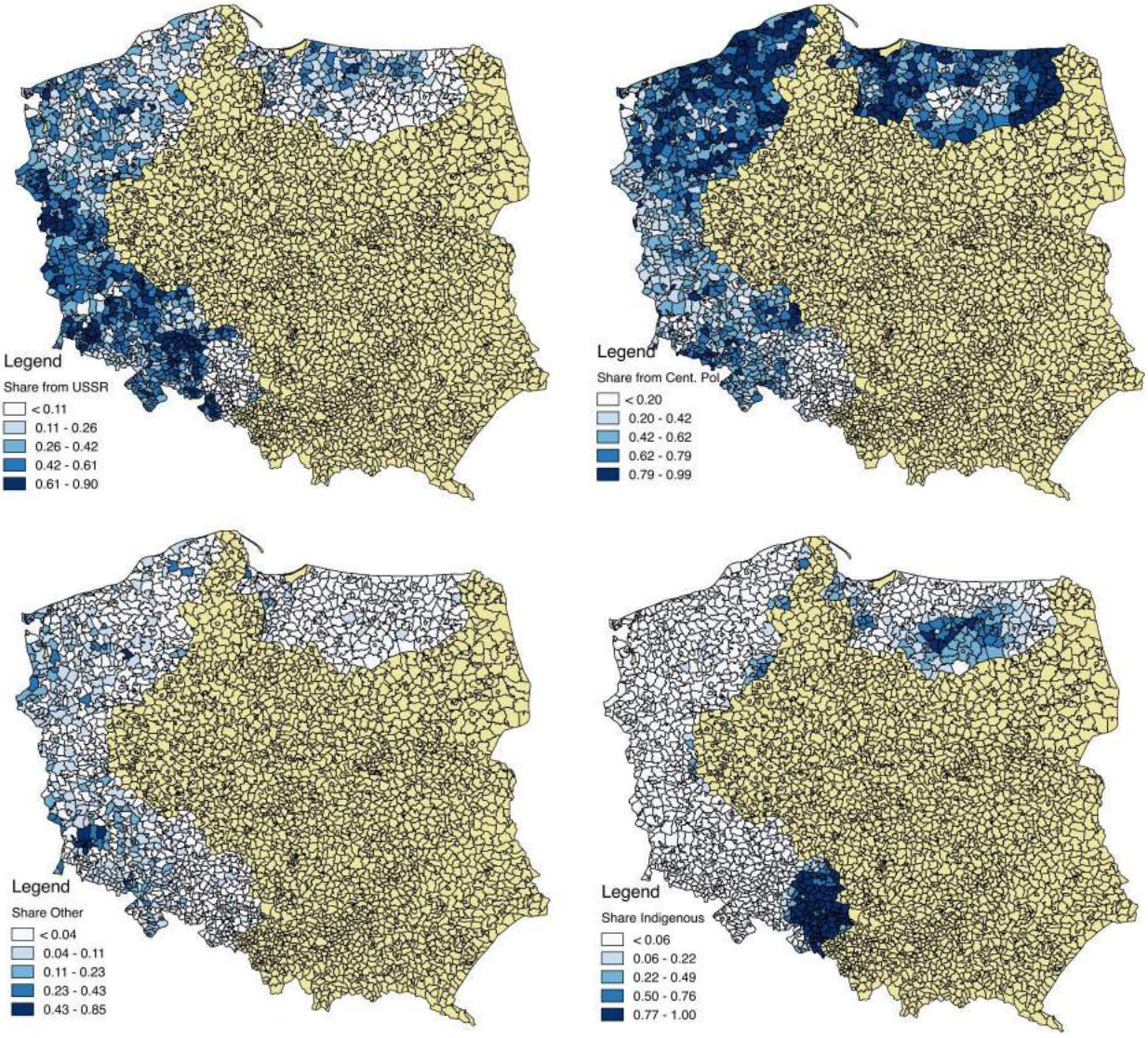
Reemigrants (*reemigranci*):

An additional, smaller group of settlers (150,000) arrived as voluntary re-emigrants from other European states (Germany, France, Belgium, Romania, Yugoslavia). They were mostly working class families, who had immigrated in the late 19th or early 20th century into the industrial centers of Europe. Those who decided to return to Poland believed they would find better career opportunities under the Communist regime. To facilitate their return, the Polish government concluded a series of agreements with the governments of states that had large Polish diasporas. The government was keen

on attracting skilled workers in the mining industry, because after the Germans had left there was a serious shortage of miners in Lower and Upper Silesia. Most of these re-emigrants, however, were manual laborers. Similarly to other groups, these settlers had little agency in deciding where to settle (Banasiak 1965).

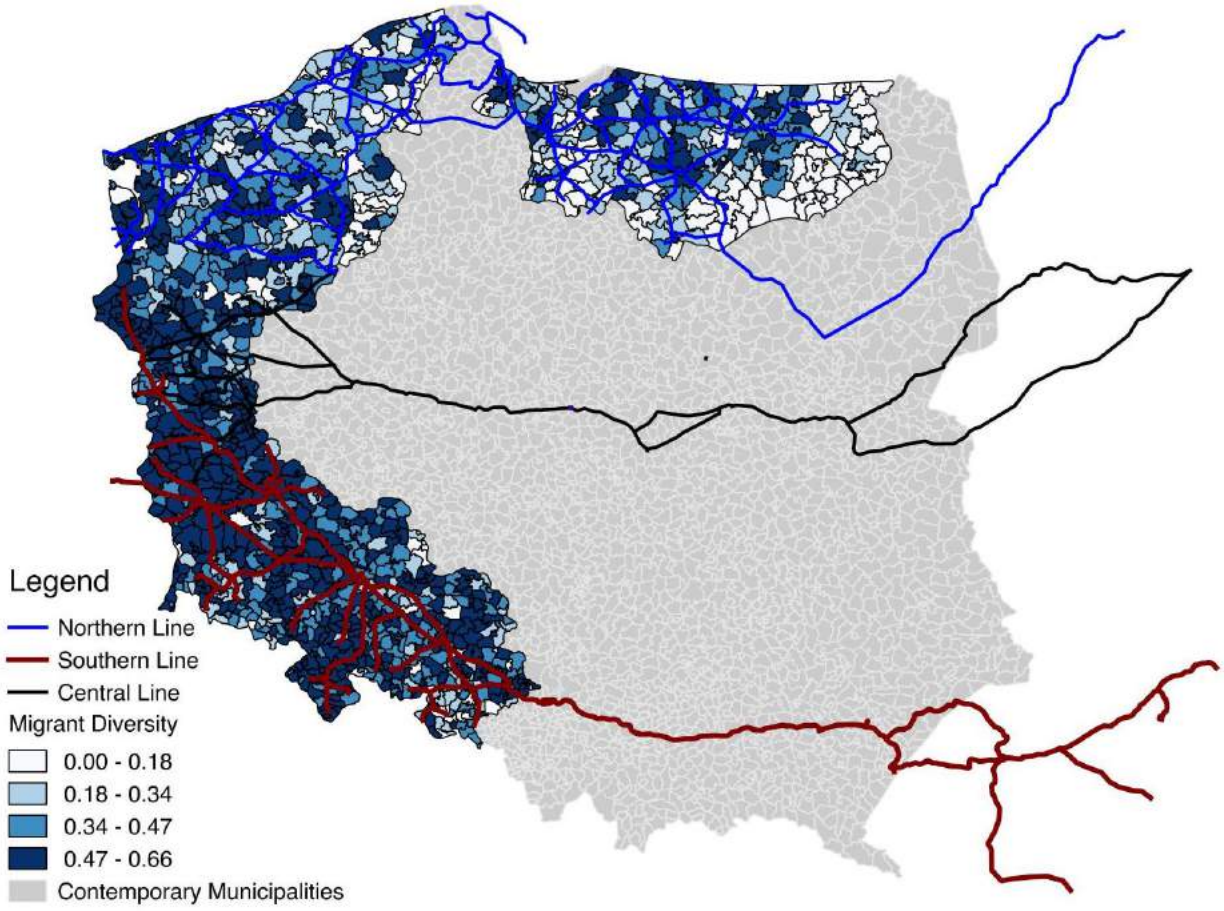
A.2 Distribution of the main population groups

Figure A1: Municipality-level Population Shares of Main Groups: (1) Repatriates from the USSR (Top Left); (2) Settlers from Central Poland (Top Right); (3) Reemigrants from Western Europe (Bottom Left); (4) Indigenous Population, including Silesians, Warmiaks, Mazurians, and Kashubians (Bottom Right).



A.3 Main railway lines used to transport migrants

Figure A2: Migrant Diversity and Primary Railway Lines Used in Population Transfers.



B Data and measurement

B.1 Sources

Municipal data on the composition of population in 1948

Archiwum Akt Nowych (AAN). Ministry for the Recovered Territories in Warsaw [1944] 1945-1949 (Ministerztwo Ziem Odzyskanych (MZO) w Warszawie [1944] 1945-1949).

- **MZO. 1515a.** Population of the Recovered Territories in 1948 – statistical tables [Ludność na Ziemiach Odzyskanych w 1948 roku – tabele statystyczne]
- **MZO. 1515j.** Population Survey on Dec. 31, 1948 in counties Elk, Goldab, Olecko – Bialystok voivodeship [Ankieta ludnościowa na 31 XII 1948 w Powiecie: Ełk, Gołdap, Olecko – województwo bialostockie]
- **MZO. 1515k.** Population Survey on Dec. 31, 1948. Status of population in Gdansk voivodeship [Ankieta ludnościowa na 31 XII 1948. Stan zaludnienia w województwie gdańskim]
- **MZO. 1515l.** Population Survey on Dec. 31, 1948. Status of population in Olstzyn voivodeship [Ankieta ludnościowa na 31 XII 1948. Ankieta ludnościowa na 31 XII 1948. Stan zaludnienia w województwie olsztyńskim]
- **MZO. 1515m.** Population Survey on Dec. 31, 1948. Status of population in Poznan voivodeship. [Ankieta ludnościowa na 31 XII 1948. Stan zaludnienia w województwie poznańskim]
- **MZO. 1515n.** Population Survey on Dec. 31, 1948. Status of population in Szczecin voivodeship. [Ankieta ludnościowa na 31 XII 1948. Stan zaludnienia w województwie szczecińskim]
- **MZO. 1515o.** Population Survey on Dec. 31, 1948. Status of population in Silesian voivodeship. [Ankieta ludnościowa na 31 XII 1948. Stan zaludnienia w województwie śląskim]
- **MZO. 1515p.** Population Survey on Dec. 31, 1948. Status of population in Wroclaw voivodeship. [Ankieta ludnościowa na 31 XII 1948. Stan zaludnienia w województwie wrocławskim]

Data from the German Census of 1939 collected by Polish MZO

Location: Archiwum Akt Nowych (AAN). Ministry for the Recovered Territories in Warsaw [1944] 1945-1949 (Ministerztwo Ziem Odzyskanych w Warszawie [1944] 1945-1949).

- **MZO. 1655.** Dane statystyczne dotyczące liczby ludności na Ziemiach Odzyskanych, stanu zatrudnienia i liczby czynnych zakładów przemysłowych. 1945-1947. B-6875. The data provided are from German Census from 17.V.1939.
- **MZO. 1656.** Powiatowe wykazy gmin na Ziemiach Odzyskanych z wyjątkiem obszaru byłego wolnego miasta Gdańska według stanu z 17.V.1939. 1945. B-6876.
- **MZO. 1957.** Powiatowe wykazy gmin na Ziemiach Odzyskanych. 1945-1946. B-6877.

Note: The 1939 Census was not conducted in the free city of Danzig and its immediate surroundings due to its special geopolitical status. MZO provided estimates for the free city of Gdansk and surrounding areas combined (without disaggregating it into communes) from other sources.

Data on distance to nearest railway station in 1948

Data for historical municipalities were published in Skorowski, Stanisław (ed.). 1948. *Podział Administracyjny Rzeczypospolitej Polskiej.* Związek Zawodowy Pracowników Samorządu Terytorialnego i Użyteczności Publicznej w Polsce. Warszawa.

Economic outcomes from the communist period

- GUS. 1950. Narodowy Spis Powszechny 1950. Warsaw: Główny Urząd Statystyczny.
- GUS. 1984. Statystyka Gmin 1983. Warsaw: Główny Urząd Statystyczny.
- GUS. 1986. Rocznik Statystyczny Miast 1985. Warsaw: Główny Urząd Statystyczny. Seria: Statystyka Regionalna.
- Roczniki Statystycze województw Katowickiego, Gdańskiego, Koszalinskiego, Ełbląskiego, Olsztynskiego, Białostockiego, Szczecińskiego, Słupskiego, Pilskiego, Zielonogorskiego, Opol-

skiego, Legnickiego, Wrocławskiego, Jeleniogórskiego, Walbrzyskiego, Gorzoskiego for the year 1981.

- Narodowy Spis Powszechny z dnia 7 XII 1978; Narodowy Spis Powszechny z dnia 16. VII. 1988.

The data on municipal guard and volunteer fire brigades (OSPs) were obtained from Project Moja Polis by Stowarzyszenie Klon/Jawor. While most volunteer fire brigades in the formerly German territories were founded in the 1940s, they were able to register following the passing of the 1989 Law on Associations in order to receive equipment, funds, and training. The majority of OSPs registered by the mid-1990s, but many organizations continued to register well into the late 1990s and the early 2000s. OSP registration data were verified against dates of founding, whenever possible. The analyses in the manuscript use information on volunteer fire brigades registered and/or verified as operating by the mid-1990s.

Data on private enterprises and personal income tax from 1995 onward are available from the website of Bank Danych Lokalnych at https://bdl.stat.gov.pl/BDL/start. Province-level data on Gross Domestic Product and Gross Income per capita in 1995 were published in GUS. 1999. Produkt Krajowy Brutto według województw za lata 1995-1997. Katowice: Główny Urząd Statystyczny.

Data on municipal taxes and personal income tax in 1993 were kindly shared by Paweł Swianiewicz of the University of Warsaw and first analyzed in Swianiewicz, Paweł. 1996. Zróżnicowanie polityk finansowych władz lokalnych. Warszawa: Inst. Badań nad Gospodarką Rynkową.

B.2 Descriptive statistics and balance

Table A1: Descriptive Statistcs for the Main Variables Used in the Analysis.

Statistic	N	Mean	St. Dev.	Min	Max
Migrant Diversity (1948)	630	0.43	0.14	0.02	0.66
Share Migrants (1948)	630	0.86	0.26	0.02	1.00
Share from USSR (1948)	630	0.28	0.19	0.002	0.84
Share from Central Poland (1948)	630	0.54	0.23	0.01	0.98
Share Autochthonous (1948)	630	0.14	0.26	0.00	0.98
Share from Europe (1948)	630	0.05	0.06	0.00	0.58
Share Men (1948)	630	0.48	0.03	0.22	0.63
Share Aged 18-59 (1948)	630	0.56	0.11	0.004	2.84
Total Population (1948)	630	8,660	17,256	360	263,104
Share Urban (1948)	630	0.22	0.32	0.00	1.00
Share in Industry (1939)	611	0.27	0.12	0.04	0.70
Share Farms Over 100 (1939)	611	0.02	0.02	0.00	0.18
Distance to Railway (1948)	630	4.22	6.64	0.00	53.80
Distance to County Seat (1950)	630	11.88	7.12	0.04	32.31
Distance to Border (km)	630	59.39	45.00	0.74	186.13
Volunteer Fire Brigades per 1,000 (1995-96)	630	4.17	4.29	0.00	21.92
Municipal Guard (2007)	630	0.27	0.44	0	1
Property Tax Rate (1993-95)	626	78.21	13.48	31.86	100.02
Property Tax Revenue per capita, Zł. (1993-95)	626	300.20	341.83	39.72	4,541.24
Employed in Socialized Economy (1982)	619	281.10	127.37	15.48	1,309.56
In Private Handicrafts per 1,000 (1982)	619	10.91	7.27	0.00	61.59
Shops per 1,000 (1980)	618	6.91	1.81	3.09	21.74
TVs per 1,000 (1980)	618	211.01	37.21	68.00	364.00
Phones per 1,000 (1980)	616	31.81	19.93	4.50	133.80
Schools per 1000 (1982)	626	5.38	6.71	0.00	90.00
Libraries per 1000 (1982)	624	3.87	4.76	0.00	79.00
Share w/ Higher Edu (1978)	594	0.02	0.02	0.00	0.23
Share w/ Higher Edu (1988)	630	0.15	0.11	0.04	1.36
Share w/ Higher Edu (2002)	629	0.06	0.03	0.003	0.26
Personal Income Tax per capita, Zł. (1993)	626	526.53	127.62	0.12	795.91
Personal Income Tax per capita, Zł. (1995)	630	101.34	21.61	8.52	154.85
Personal Income Tax per capita, Zł. (1998)	630	169.12	65.13	13.60	656.68
Personal Income Tax per capita, Zł. (2000)	630	143.85	50.91	8.44	421.57
Private Enterprises per 1,000 (1995)	630	36.66	20.54	5.82	214.36
Private Enterprises per 1000 (1998)	630	54.51	25.80	2.66	276.77
Private Enterprises per 1000 (2000)	630	63.15	28.96	3.88	340.45

Table A2: Relationship between *Migrant Diversity* and Socio-Economic Covariates as well as the Share of Migrants at the Level of Contemporary Municipalities.

Variable	**Pearson Correlation**	**R^2 (variance explained)**
Share Migrants (1948)	-0.01	0.00
Share Male (1948)	-0.01	0.00
Share Aged 18-59 (1948)	0.00	0.00
Share Urban (1948)	0.08	0.01
Share in Industry (1939)	0.34	0.12
Share Farms over 100 ha (1939)	-0.08	0.01
Ln population (1948)	0.14	0.02
Distance to county seat (1950)	-0.12	0.01
Ln Distance to County Seat (1950)	-0.09	0.01
Distance to Raiway (1948)	-0.29	0.09
Distance to Border	-0.33	0.11
Ln Distance to Border	-0.25	0.06

B.3 Unit of analysis

The analysis is performed at the level of municipality (*gmina*), the smallest unit for which information on socio-economic outcomes in the 1980s and 1990s is available. Municipalities are self-contained social units with legislative and governing bodies and thus appropriate for studying the relationship between the cultural composition of a community, variation in enforcement mechanisms, and economic activity.

However, municipal boundaries changed considerably during the period under study. In the 1948 census, towns and rural municipalities, comprising a group of closely situated villages, were listed separately. In 1954, municipalities were reorganized into even smaller units (*gromady*), and in 1973 gromady were abolished and municipalities (*gminy*) were reintroduced with different borders than the pre-1954 units. Between 1973 and 2013, municipality boundaries underwent further changes. In particular, some small towns were joined to the neighboring rural municipalities and recategorized into "urban-rural" municipalities. Towns were classified as "urban" municipalities and groups of villages as "rural" municipalities.

Thus, although "treatment" of cultural diversity is measured at the level of smaller, 1948 municipalities, I conduct analysis at the level of contemporary, larger municipalities, to account for the various administrative changes. To match units from different time periods to each other, I digitized and georeferenced the map of municipalities printed by the Central Office for National Measurements (*Główny Urzad Pomiarów Kraju*) for internal use in 1949. I then superimposed this map onto a shapefile of the contemporary Polish municipalities and assigned each historical unit to the contemporary unit that covered most of its territory (see Figure A3). Both procedures were conducted in ArcGIS. Where contemporary municipality borders split historical municipalities, I weighted the historical data by the proportion of the overlapping area. This method assumes homogeneous distribution of population across territory. While this is an oversimplication, it results in relatively low distortion due to the small area of municipalities. As a result, 1,217 historical municipalities mapped onto approximately 630 contemporary units.[1]

[1]Historical data could not be extrapolated for municipalities Rewal, Mielno, and Krynica Morska due to the shape of their borders and location.

Fortunately, the pre-treatment covariates from the 1939 German census are available at the level of even smaller units, communes (*gemeinde*), which are roughly equivalent to Polish villages or towns without surrounding villages and as a result were not split across municipal boundaries established in post-1945 period. However, the 1939 data do not cover the Danzig/Gdansk and the surrounding areas.

I also use dummies for German districts (*Regierungsbezirke*). These are Liegnitz, Oppeln, and Breslau in Silesia; Königsberg, Allenstein, Gumbinnen, and Marienwerder (Westpreussen); Frankfurt and Potsdam in Brandenburg; and Stettin and Köslin in Pomerania.

Figure A3: Historical Units (Red, Depicted only for the Formerly German Territories) versus Contemporary Units (Blue).

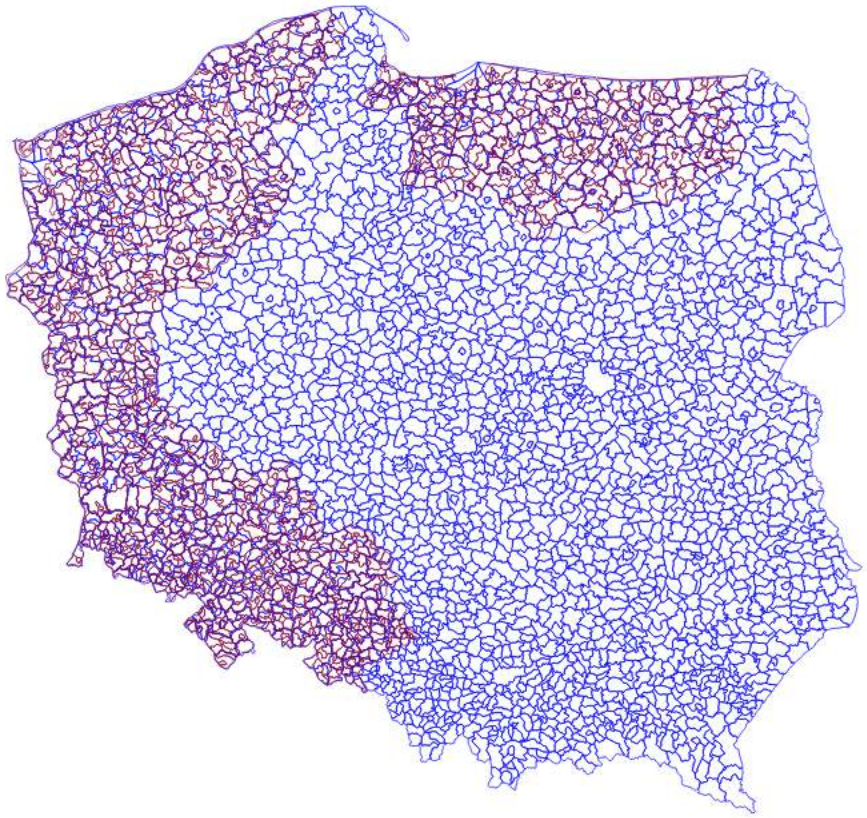
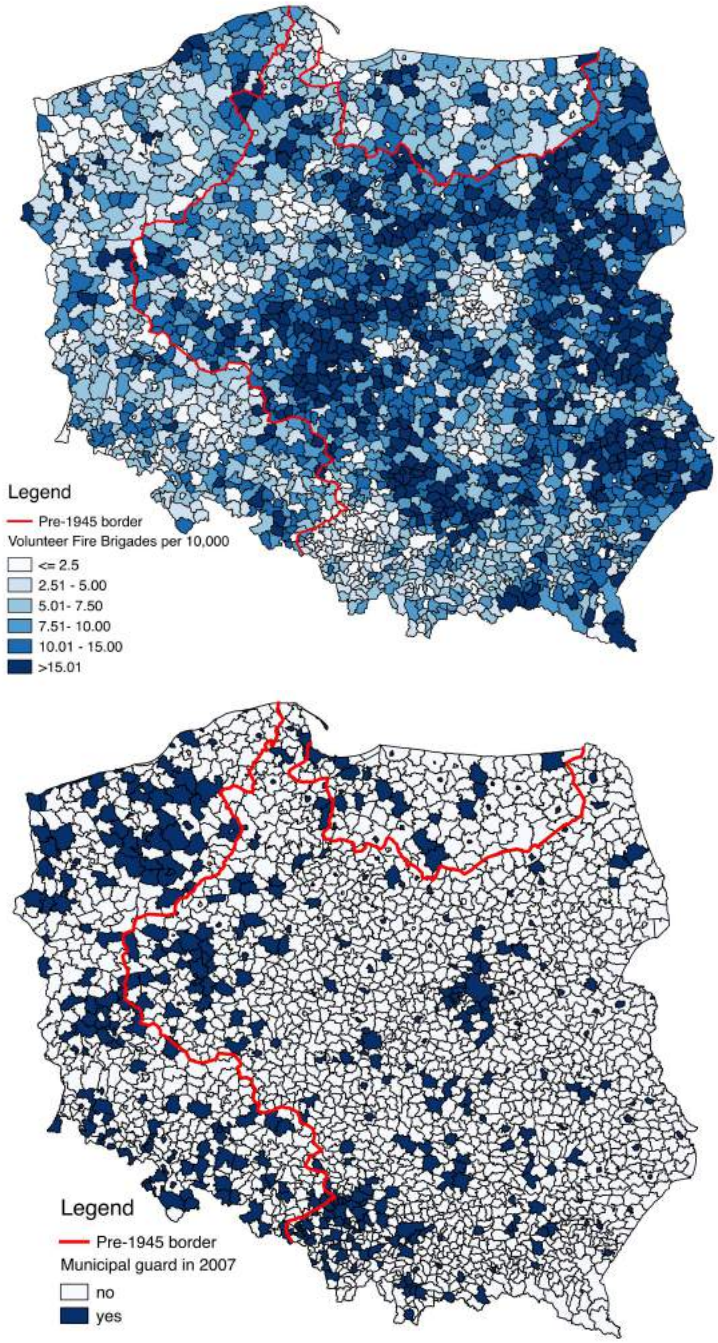


Figure A4: Presence of Volunteer fire Brigades and Municipal Guard Units (*Straz Gminna*) in Contemporary Poland.



C Resettled territories during the communist period

C.1 Membership in the Communist Party

The shift of Poland's borders and subsequent population transfers contributed to the growth of Communist influence. In the resettled territories, the party-state was able to penetrate and reorganize parts of society, blurring the distinction between public and private. It became autonomous not only from the social forces but also from the constraints imposed by its own laws, which could be changed and/or enforced at will.

On December 15, 1948, the PPR and the Polish Socialist Party (*Polska Partia Socjalistyczna, PPS*) were joined into the Polish United Workers' Party (*Polska Zjednoczona Partia Robotnicza, PZPR*). At that time, Poland became a de facto single-party state, with PZPR and allied political forces retaining power until 1989. The unification was also an opportunity to count membership at the county-level. Figure A5 plots the number of party members (divided by population). The map shows that membership was higher in the formerly German territories than in other parts of Poland. Figure A6 shows that Communist Party membership increases with heterogeneity.

Figure A5: PZPR Membership in December 1948.

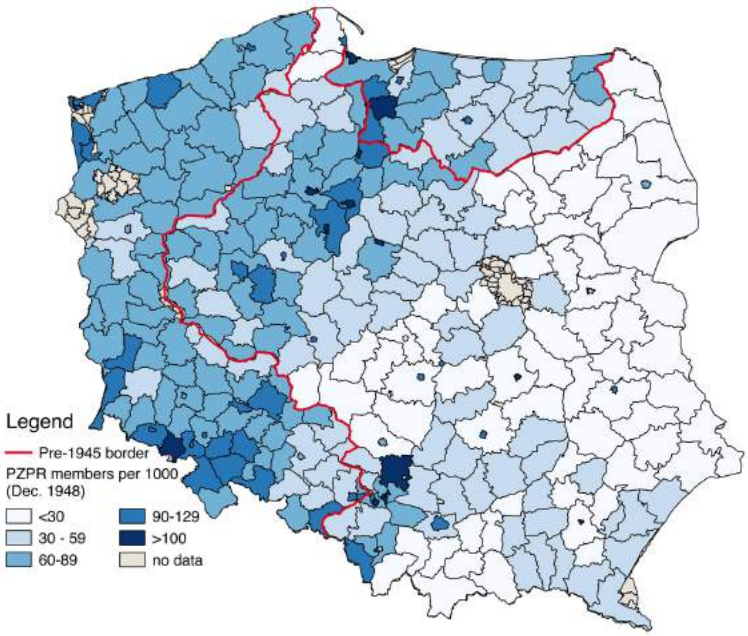
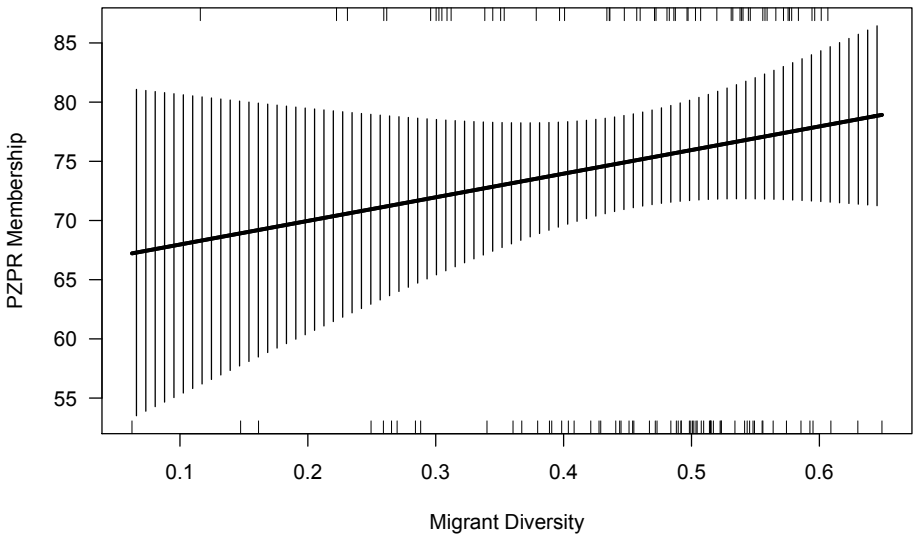


Figure A6: Migrant Diversity and PZPR Membership at the end of 1948.



C.2 Communist economic policies

Table A3: Stalinism in the Formerly German Territories and in Other Parts of Poland.

	Resettled provinces	**Non-Resettled provinces**
State agricultural farms (PGR) in 1955 *per 10,000 people*	5.82	1.10
Output of private industry in 1955 *per person, in Zloty*	3.52	10.91
Output of state industry in 1955 *per person, in Zloty*	1451	1571
Private Crafts Workshops in 1957 *per 1000 people*	4.49	4.88

Sources: GUS. 1956. Rocznik Statystyczny 1956. Warsaw: Główny Urząd Statystyczny. GUS. 1958. Rocznik Statystyczny 1958. Warsaw: Główny Urząd Statystyczny.

C.3 Crime rates in the 1960s

County-level data on crimes in the formerly German territories are available from Province Statistical Yearbooks (*Roczniki Statystyczne Województw*). The earliest data covering provinces affected by population transfers (Wrocław, Olsztyn, Opole, Gdańsk, Koszalin, and Białystok) go back to 1961-1962. Crime rates are measured as the number of crimes per 1,000 people. I control for the share of urban population (*Share Urban*), share of the population in Industry (*Share in Industry*), location next to the German border (*Border* dummy), the share of large farms (*Farms over 100 ha*). Because the data were published in Province Yearbooks and may reflect differences in categorization and collection of statistical information, I include province-fixed effects rather than dummies for German districts.

Regression analysis in Table A4 demonstrates that *Migrant Diversity* predicts higher crime rates, including higher prevalence of robberies and hooliganism. The magnitude of the coefficient on *Migrant Diversity* in Model 1 suggests that most heterogeneous counties had 6.56 more crimes per 1,000 people than homogeneous counties, equivalent to one standard deviation in the dependent variable. The coefficient on *Migrant Diversity* is positive but does not reach conventional levels of statistical significance in Model 2, which isolates the rates of thefts and burglaries and is positive and statistically significant at the 5% level for robberies and hooliganism. Higher prevalence of hooliganism, typically by young people in urban areas, speaks to lower levels of informal social control and/or potentially greater willingness to report this relatively minor crime to the state police in more culturally heterogeneous localities.[2]

[2]The data on hooliganism are not available for Olsztyn province.

Table A4: Crime Rates (per 1,000 people) in the Resettled Counties in Wrocław, Koszalin, Opole, Gdansk, Białystok, and Olsztyn Provinces in 1961-1962. OLS regression.

	Crime rates per 1000 people			
	All Crimes	Theft	Robbery	Hooliganism
	(1)	(2)	(3)	(4)
Migrant Diversity	12.33^{+}	7.42	0.23*	2.89*
	(7.35)	(5.03)	(0.11)	(1.33)
Share Migrants	4.81	1.77	0.03	0.33
	(3.44)	(2.36)	(0.05)	(0.55)
Covariates	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
N	91	91	91	74
Adjusted R^2	0.38	0.24	0.20	0.28

$^{+}p < .1$; $^{**}p < .05$; $^{***}p < .01$

D Additional results and robustness checks

D.1 Alternative regression specifications

Table A5: Models omitting *Share Migrants*, the non-exogenous component of population heterogeneity in the resettled territories.

	Volunteer Fire Brigades		Municipal Guard		ln(Property Tax Revenue)	
	OLS		logistic		OLS	
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant Diversity	-3.75^{**}	-3.91^{**}	1.90^{**}	2.20^{+}	1.20^{**}	0.67^{**}
	(1.24)	(1.47)	(0.71)	(1.20)	(0.19)	(0.22)
Covariates		✓		✓		✓
District FE		✓		✓		✓
N	630	611	630	611	626	607
Adjusted R^2	0.01	0.24			0.06	0.24
Log Likelihood			−362.56	−242.34		
AIC			729.13	522.69		

+ p<0.1; * p<0.05; ** p<0.01

Table A6: Models 1, 3, 4, and 6 omit *Share Migrants*, the non-exogenous component of population heterogeneity in the resettled territories. Models 2 and 5 exclude most covariates.

	ln(Telephones per 1000)			ln(Shops per 1000)		
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant Diversity	-0.33^{**}	-0.33^{**}	-0.18^{*}	−0.27	−0.26	-0.39^{**}
	(0.07)	(0.07)	(0.08)	(0.17)	(0.16)	(0.15)
Share Migrants		0.11^{**}			0.63^{**}	
		(0.04)			(0.08)	
Covariates			✓			✓
District FE			✓			✓
N	618	618	601	616	616	599
Adjusted R^2	0.03	0.05	0.25	0.003	0.09	0.54

+ p<0.1; * p<0.05; ** p<0.01

Table A7: Models omitting *Share Migrants*, the non-exogenous component of population heterogeneity in the resettled territories (Continued).

	ln(Personal Income Tax per capita)				ln(Private Enterprises per 1,000)			
	1995		1998		1995		1998	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Migrant Diversity	0.34**	0.23**	0.48**	0.25**	0.70**	0.30*	0.56**	0.22*
	(0.06)	(0.08)	(0.08)	(0.08)	(0.14)	(0.14)	(0.12)	(0.11)
Covariates		✓		✓		✓		✓
District FE		✓		✓		✓		✓
N	630	611	630	611	630	611	630	611
Adjusted R^2	0.04	0.24	0.05	0.35	0.04	0.47	0.03	0.51

+ p<0.1; * p<0.05; ** p<0.01

Table A8: *Migrant Diversity* and economic outcomes in 1995, 1998, and 2000. Dependent variables are (1) personal *Income Tax per Capita* (in *Złoty*) in Models 1-3 and (2) *Private Enterprises* per 1,000 people in Models 4-6. OLS regression, minimum specifications.

	ln(Personal Income Tax)			ln(Private Enterprises)		
	1995	1998	2000	1995	1998	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant Diversity	0.34**	0.48**	0.73**	0.71**	0.57**	0.46**
	(0.06)	(0.08)	(0.10)	(0.14)	(0.12)	(0.11)
Share Migrants	−0.12**	−0.10*	0.01	0.28**	0.36**	0.35**
	(0.03)	(0.04)	(0.05)	(0.07)	(0.06)	(0.06)
N	630	630	630	630	630	630
Adjusted R^2	0.06	0.06	0.08	0.06	0.08	0.08

+ p<0.1; * p<0.05; ** p<0.01

Table A9: *Migrant Diversity*, Personal Income Tax, and Private Entrepreneurship in 1995, 1998, and 2000. OLS Regression, fully-specified models without spatial filtering.

	ln(Personal Income Tax per capita)			ln(Private Enterprises per 1,000)		
	1995	1998	2000	1995	1998	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant Diversity	0.21**	0.24**	0.27**	0.30*	0.23*	0.22*
	(0.08)	(0.09)	(0.09)	(0.14)	(0.11)	(0.10)
Share Migrants	−0.20**	−0.09	0.04	−0.05	0.06	0.10
	(0.06)	(0.06)	(0.06)	(0.10)	(0.08)	(0.08)
Share Urban	−0.06+	0.14**	0.39**	0.60**	0.55**	0.57**
	(0.03)	(0.04)	(0.04)	(0.06)	(0.05)	(0.04)
Share in Industry	0.03	0.45**	0.80**	0.98**	0.75**	0.58**
	(0.10)	(0.11)	(0.11)	(0.17)	(0.14)	(0.13)
Share Farms over 100 ha	1.29*	0.79	−0.63	−2.18*	−2.85**	−2.23**
	(0.50)	(0.56)	(0.58)	(0.92)	(0.72)	(0.68)
ln(Population)	0.04*	0.12**	0.05**	−0.02	−0.04+	−0.04*
	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Distance to State Services	0.002	−0.0004	−0.004*	−0.01**	−0.01**	−0.01**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Distance to Border	−0.0001	0.0004	0.0000	0.002**	−0.0002	−0.0004
	(0.0003)	(0.0003)	(0.0003)	(0.001)	(0.0004)	(0.0004)
Distance to Railway	−0.005**	−0.003+	−0.002	−0.01*	−0.002	−0.001
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
District FE	✓	✓	✓	✓	✓	✓
N	611	611	611	611	611	611
Adjusted R^2	0.26	0.35	0.52	0.47	0.51	0.51

+ p<0.1; * p<0.05; ** p<0.01

Table A10: Diversity and Economic Outcomes in 1995, 1998, and 2000 in a sample *Without Cities*. Dependent Variables are Personal Income Tax (in *Złoty*), Models 1-3; Private Enterprises, Models 4-6. OLS Regression.

	ln(Personal Income Tax per capita)			ln(Private Enterprises per 1,000)		
	1995	1998	2000	1995	1998	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant Diversity	0.25**	0.25**	0.30**	0.41**	0.29*	0.28*
	(0.07)	(0.07)	(0.08)	(0.15)	(0.12)	(0.11)
Share Migrants	−0.20**	−0.10*	0.05	−0.01	0.08	0.12
	(0.05)	(0.05)	(0.05)	(0.10)	(0.08)	(0.08)
Covariates	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
N	549	549	549	549	549	549
Adjusted R^2	0.34	0.38	0.51	0.36	0.41	0.40

+ p<0.1; * p<0.05; ** p<0.01

Table A11: Voluntary Provision of Public Goods, Tax Revenues (averaged for 1993-95), and Personal Income Tax and Private Enterprises (2000). OLS Regression.

	ln(Personal Income Tax per capita)			ln(Private Enterprises per 1,000)		
	(1)	(2)	(3)	(4)	(5)	(6)
Volunteer Fire Brigades	-0.01^{**}			-0.01^{*}		
	(0.002)			(0.003)		
Tax Rate		0.004^{**}			0.005^{**}	
		(0.001)			(0.001)	
log(Tax Revenue)			0.15^{**}			0.20^{**}
			(0.02)			(0.02)
Covariates	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
N	611	607	607	611	607	607
Adjusted R^2	0.53	0.54	0.58	0.51	0.53	0.60

+ p<0.1; * p<0.05; ** p<0.01

Table A12: Relationship between per Capita Capital Spending (1993-95) and Migrant Diversity, Property Tax Rates and Revenues. OLS Regression.

	ln(Capital Spending per capita)			
	(1)	(2)	(3)	(4)
ln(Property Tax Revenue)	0.34^{**}	0.41^{**}		
	(0.05)	(0.06)		
Property Tax Rate			0.01^{**}	0.01^{*}
			(0.003)	(0.003)
Covariates		✓		✓
District FE		✓		✓
N	626	607	626	607
Adjusted R^2	0.06	0.09	0.01	0.04

+ p<0.1; * p<0.05; ** p<0.01

D.2 Non-parametric CBGPS weighting

Analysis in this section relies on non-parametric covariate balancing generalized propensity score weighting (npCBGPS) methodology developed by Fong et al. (2018). Two main advantages of this approach are its robustness to the misspecification of the propensity score model and its applicability to continuous treatment variables (Fong, Hazlett, and Imai 2018). Non-parametric CBGPS produces weights that can be used as regression adjustment, minimizing the association between the treatment and observed covariates and ensuring that the differences in outcomes are due to treatment rather than pre-treatment differences.

The npCBGPS approach assumes that the treatment (*Migrant Diversity*) is normally distributed. To make this assumption more plausible, I transform the variable using formula

$$-(max(X)+1)/\sqrt{max(X)+1-X}$$

The pre-treatment covariates in the generalized propensity score model include all covariates used in the OLS regressions in the article as well as the gender and age of migrants available from the 1948 census. I also add the squares of non-binary covariates to balance both first and second moments. Table A13 shows that npCBGPS substantially improves balance, resulting in near-zero correlations between treatment and the covariates.

The outcome models include the treatment variable on original scale (*Migrant Diversity*) and fixed effects for German districts as well as weights obtained from the npCBGPS. In Model 1 Table A14 the coefficient on *Volunteer Fire Brigades* increases in magnitude relative to the fully-specified model in the article. The coefficient on *Tax Share* increases in magnitute relative to the OLS models without weighting and is now significant at a 5% level. The coefficient on *Tax Revenue* is broadly similar to that in OLS without weighting presented in Table 2. The coefficient on *Migrant Diversity* increases in magnitude and remains statistically significant for *Shops*, but is smaller and no longer significant for *Phones*. Finally, the coefficients on *Migrant Diversity* increase in magnitude and retain statistical significance for all post-1989 economic outcomes (Models 1-6 in Table A15).

Table A13: Pearson Correlation between transformed and non-transformed *Migrant Diversity* and first and second moments of demographic and economic covariates.

Covariate	Unweighted		Weighted
	No transf.	Transf.	Transf.
Share Migrants (1948)	-0.01	-0.03	-0.03
Share Men (1948)	0.00	0.00	-0.01
Share Aged 18-59	0.01	0.01	-0.04
Share Urban (1948)	0.08	0.06	-0.00
Share in Industry (1939)	0.34	0.35	-0.03
Share Farms over 100 ha (1939)	-0.08	-0.10	-0.01
ln(Population) (1948)	0.14	0.13	0.02
Distance to County Seat (1950)	-0.11	-0.10	-0.02
Distance to Border	-0.34	-0.34	0.00
Distance to Railway (1948)	-0.32	-0.30	0.00
Share Migrants ^2	0.05	0.02	-0.03
Share men ^ 2	0.00	-0.00	-0.02
Share aged 18-59 ^2	-0.06	-0.05	-0.04
Share Urban ^2	0.05	0.03	-0.00
Share in Industry ^2	0.30	0.32	-0.03
Share Farms over 100 ha ^2	-0.06	-0.07	-0.01
ln(Population)^2	0.14	0.13	0.02
Distance to County Seat ^2	-0.10	-0.09	-0.02
Distance to Border ^2	-0.34	-0.34	-0.00
Distance to Railway ^2	-0.30	-0.27	-0.02

Table A14: OLS regression with npGBGPS weighting. Dependent variables are Volunteer Fire Brigades per 1,000 people (Model 1); Property Tax Share in 1993-95 (Model 2); and Property Tax Per Capita (Model 3); Shops (Public and Private) (Model 4); and Phones per 1,000 people (Model 5).

	Fire Brigades	Tax Share	ln(Tax Revenue)	ln(Shops)	ln(Phones)
	(1)	(2)	(3)	(4)	(5)
(6)					
Migrant Diversity	-4.91^{**}	11.92^{*}	0.59^{**}	-0.26^{**}	-0.35
	(1.53)	(5.13)	(0.22)	(0.08)	(0.21)
District FE	✓	✓	✓	✓	✓
N	611	607	607	601	599
Adjusted R^2	0.07	0.06	0.07	0.08	0.08

+ p<0.1; * p<0.05; ** p<0.01

Table A15: Diversity and Post-Communist Economic Outcomes. Dependent variables are*Personal Income Tax* (Models 1-3) and *Private Enterprises* (Models 4-6). OLS regression with npGBGPS weighting.

	ln(Personal Income Tax per capita)			ln(Private Enterprises per 1,000 people)		
	1995	1998	2000	1995	1998	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant Diversity	0.25**	0.30**	0.54**	0.35*	0.25+	0.23+
	(0.07)	(0.09)	(0.11)	(0.17)	(0.14)	(0.13)
District FE	✓	✓	✓	✓	✓	✓
N	611	611	611	611	611	611
Adjusted R^2	0.33	0.09	0.10	0.15	0.16	0.15

+ p<0.1; * p<0.05; ** p<0.01

D.3 Sensitivity to unmeasured confounders

I explore the sensitivity of the main estimates to the violation of the assumption of no unmeasured confounders using the graphic tools developed Carnegie et al.(2016) and implemented in R package *treatSens.* Unlike most other existing approaches to sensitivity analysis,[3] *treatSens* works well for continuous treatments, does not require matched samples, and is intuitive to interpret.[4]

TreatSens characterizes the unmeasured confounder using two sensitivity parameters: (1) partial correlation between an unobserved confounder and treatment assignment and (2) partial correlation between an unobserved confounder and outcome (Carnegie, Harada, and Hill 2016). The results of sensitivity analysis are summarized in a contour plot. Each black contour line represents the combinations of sensitivity parameters for the unobserved confounder that lead to the same estimated treatment effect. The red line labeled 0 represents a combination of sensitivity parameters that a single confounder would need to have in order to drive the treatment estimate to zero; the blue lines represent the region in which treatment effect estimate is not significant at the 5% level. The x- and y-axes represent zero confounding. The observed treatment effect estimate from the OLS model is on the x-axis (Carnegie, Harada, and Hill 2016). The coefficient estimates on observed covariates, scaled to have mean of zero and standard deviation of one, can be used as a benchmark for interpretation. In a contour plot these are depicted by pluses (for observed confounders with positive partial association with the outcome) and inverted triangles (for observed confounders with negative partial association with the outcome, which are reverse-scaled) (Carnegie, Harada, and Hill 2016).

Results of sensitivity analyses are presented graphically on the next few pages.[5] The contour plot in Figure A7, *DV: Volunteer Fire Brigades*, shows that over a range of plausible sensitivity parameters, the estimated treatment effect is statistically significant and negative. The red contour line indicates that to attenuate the coefficient estimate on *Migrant Diversity* to zero, the unobserved confounder would have to be negatively correlated with the treatment (*Migrant Diversity*) and positively correlated

[3]For example, sensitivity tools developed by Rosenbaum (2002) require matched pairs of treated and control units.

[4]The downside is that this approach can be used only for continuous outcomes, so I do not explore the sensitivity of the results for *Municipal Guards* to unobserved confounding.

[5]Prior to sensitivity analysis, I standardized all continuous covariates.District fixed effects and moran eigenvectors are omitted because they prevented models from converging.

with the response variable (*Volunteer Fire Brigades*) and, more important, significantly more powerful in predicting the outcome than most of the observed confounders, as indicated by the distance of pluses and inverted triangles from the red line (0 effect size). To raise the p-value above 0.05, an unobserved confounder would have to be similar in predictive power to that of observed confounders depicted as an inverted triangle on the blue line (*Share urban*). Failing to observe a confounder of such magnitude is possible but unlikely, as I already account for the main factors that influence the creation of volunteer fire brigades (industry, urbanization, population size, and distance to state services). Thus, the sensitivity analysis increases our confidence in the finding that *Migrant Diversity* negatively affected the creation of *Volunteer Fire Brigades* in Polish villages.

The two plots in Figure A8 focus on two economic outcomes during state socialism for which regressions in the paper suggested a statistically significant negative effect of *Migrant Heterogeneity*. The plots for both *DV: Shops per 1000, 1982* and *DV: Phones per 1000, 1982* indicate that under all plausible sensitivity parameters, most of the range for the estimated treatment effects are negative and statistically significant. However, the results for *Shops* indicate that an unobserved confounder with predictive power comparable to any of the three observed confounders (*Share urban*, *Distance to State Services*, and *Distance to Border*) can reduce this effect to zero. For *Phones* there is only one observed confounder with equivalent statistical power (*Share Urban*).

Figure A7: Sensitivity to unobserved confounding for *Volunteer Fire Brigades* per 1,000 people.

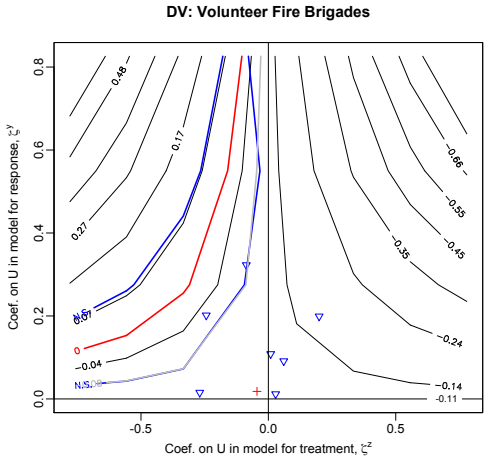
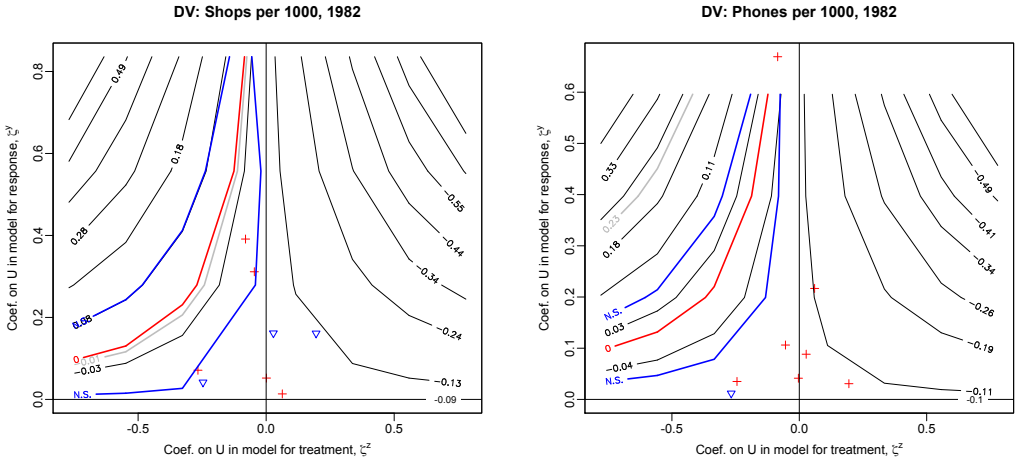
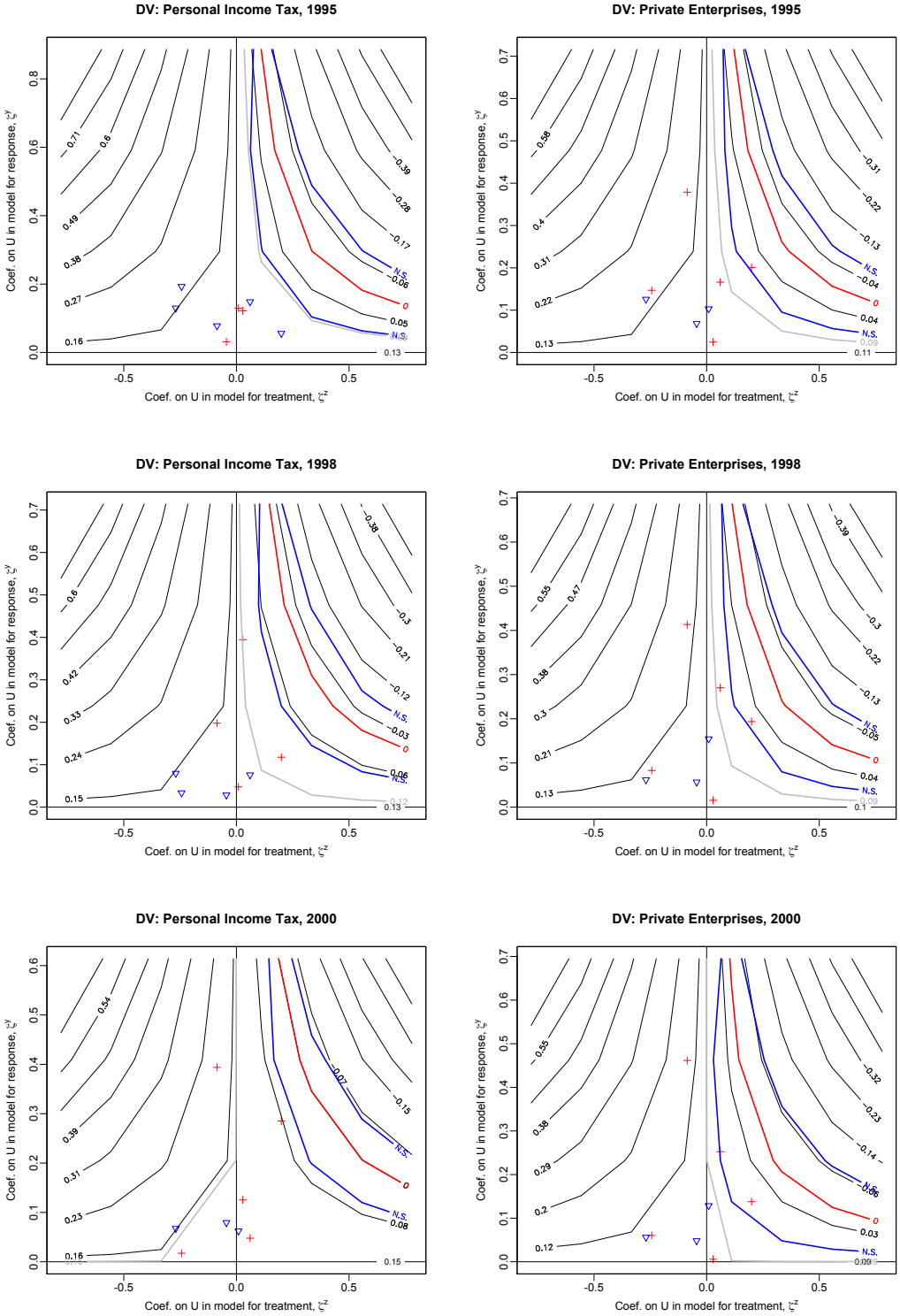


Figure A8: Sensitivity to unobserved confounding for (1) *Shops* per 1000 people in 1982; and (2) Phones per 1000 people in 1982.



The final set of contour plots in Figure A9 examines the sensitivity of the effect of *Migrant Diversity* on post-1989 economic outcomes to unobserved confounders. Plots on the left focus on *Personal Income Tax* in 1995, 1998, and 2000. The range for the treatment effect is largely positive and statistically significant. Again, the predictive power of an unobserved confounder would have to exceed the power of any of the observed confounders to reduce the treatment effect to zero or to raise the p-value above 0.05. Thus, sensitivity analysis strengthens the confidence that *Migrant Diversity* has a positive and statistically significant effect on the size of *Personal Income Tax* after 1989. Plots on the right in Figure A9 focus on the relationship between *Migrant Diversity* and the prevalence of *Private Enterprises* in post-1989 period. Although most of the range of the estimated treatment effect is still positive for all plausible sensitivity parameters, one observed confounder (*Share in Industry*) falls within the $\alpha = 0.05$ statistical significance cut-off. For 2000, *Share Migrants* is also touching the blue line. This means that an unobserved confounder of comparable predictive power can raise the p-values on the *Migrant Diversity* coefficient beyond 0.05. However, such an unobserved confounder would also have to exceed the predictive power of all of the remaining observed confounders. Since the models already control for key predictors of economic activity, including urbanization, industrialization, and railway infrastructure and have high R^2 after correcting for spatial dependence, this is not very likely.

Figure A9: Sensitivity to unobserved confounding for economic activity in post-1989 Poland: (1) *Personal Income Tax* in 1995, 1998, and 2000; (2) *Private Enterprises* per 1,000 people in 1995, 1998, and 2000.



D.4 Sequential g-estimation

I argue that contemporary economic differences between communities at different levels of heterogeneity are due to the differences in state-society relationship. I thus theorize a mediation relationship, whereby historical levels of diversity affect post-communist economic development via reliance on informal/formal institutions and the resulting levels of state capacity. Testing the mediation relationship requires more assumptions than simple regression analysis presented in Tables 2, 3, and 4 in the article, which shows that *Migrant Diversity* predicts the reliance on informal enforcement for the provision of public goods, willingness and ability to tax, and private economic activity and incomes.

One way to test whether institutions indeed mediate the effects of *Migrant Diversity* on economic outcomes is sequential g-estimation approach that originated in biostatistics literature and has been applied in political science contexts in Acharya, Blackwell, and Sen (2016). Sequential g-estimation allows estimating the controlled direct effect of heterogeneity with the hypothesized mediator fixed at a specific value for all units and thus avoiding intermediate-variable bias that arises from simply adding both the mediator and the treatment variables into the model. Intermediate-variable bias is a type of selection bias that arises due to the presence of intermediate confounders (variables affected by treatment that also affect the mediator and the outcome) (Acharya, Blackwell, and Sen 2016, 2).

The sequential g-estimation relies on two assumptions, which are quite restrictive but still less demanding than those necessary for other mediation approaches. The first is sequential unconfoundedness, or (1) no omitted variables for the effect of treatment on outcome, conditional on pre-treatment outcomes, and (2) no omitted variables for the effect of the mediator on outcome, conditional on treatment, pretreatment confounders, and intermediate confounders (Acharya, Blackwell, and Sen 2016, 8). Sequential uncounfoundedness is violated if there are omitted variables for the relationship between mediator (voluntary public goods provision, fiscal capacity) and outcome (e.g., entrepreneurship rates, income). The second assumption, no intermediate interactions, requires the effect of mediator on the outcome to be independent of the intermediate confounders. If this assumption is violated, sequential g-estimation will estimate weighted averages of average controlled direct effect of treatment within levels of intermediate confounders, i.e., this assumption is less restrictive than sequential

unconfoundedness.

Here I focus on post-communist economic outcomes because of limited data availability for the period between the earliest community-level measures of economic activity (1980-82) and the completion of the resettlement (1948). I look at two related mediator variables: (1) voluntary provision of local public goods, proxied by the density of volunteer fire brigades, and (2) fiscal capacity, proxied by property tax revenues.

To satisfy the sequential unconfoundedness assumption, the first stage of the sequential g-estimation includes not only the pretreatment confounders used in the main analysis, but also the following intermediate confounders from the communist period: the number of schools and state farms; the share of the population employed in the socialized sector, the share of the population in private handicrafts, and the number of Shops, TV-sets and Phones per 1,000 people.[6] The results of the first-stage analysis used to estimate the demediation function are presented in Table A16 in odd-numbered models.

The second stage of the sequential g-estimation uses the results from the first stage to de-mediate the outcome, subtracting the variation caused by the mediator. This stage includes only the pretreatment covariates. The coefficient on the treatment variable from this stage is interpreted as an estimate of the variation in the outcome due to the direct effect of treatment (*Migrant Diversity*) that does *not* travel through the hypothesized mediators. Results from this stage are presented in the even-numbered models of Table A16.

In models that use the demediated personal income tax per capita as the outcome, the coefficient on *Migrant Diversity* decreases in magnitude, but remains positive and statistically significant. This suggests that, in addition to the effect through institutions, *Migrant Diversity* may also have a direct effect on income tax per capita. In other words, institutional differences between communities at different levels of heterogeneity alone cannot explain the post-communist divergence in per capita income tax. One possibility is that differences in income tax at the community level result from the combination of higher levels of private economic activity and greater skill complementarity in more heterogeneous communities and thus combine a direct and an indirect effects of *Migrant Diversity*.

An important caveat is that these estimates might still be biased if there remain unmeasured con-

[6]The last five variables were used as outcomes in the analysis of communist economic outcomes.

founders for the relationship between reliance on informal institutions and economic outcomes. Figure A10 presents the results of the sensitivity analyses for the first mediator, *Volunteer Fire Brigades*, graphically. The x-axis represents the residual correlation between the mediator and outcome after accounting for the observed baseline and intermediate confounders (i.e., the amount of unmeasured confounding); the y-axis presents the estimate of the effect of *Migrant Diversity* on *Entrepreneurship* for a given correlation. We see that the estimated controlled direct effect of *Migrant Diversity* on *Private Enterprises* is indistinguisheable from zero for all positive correlations, a wide interval. This increases confidence in the conclusion that institutional differences between heterogeneous and homogeneous communities mediate the relationship between heterogeneity and post-communist economic outcomes. By contrast, the estimated controlled direct effect of *Migrant Diversity* on *Personal Income Tax* is positive unless the correlation between the mediator and outcome errors is greater than 0.25. This is not very likely as participation in volunteer fire brigades does not increase personal incomes. Sensitivity analysis thus indicates that *Migrant Diversity* has a positive effect on income that does not travel through institutional channels.

Figure A10: Sensitivity analysis for volunteer fire brigades as the mediator. Shaded regions are 95% confidence intervals.

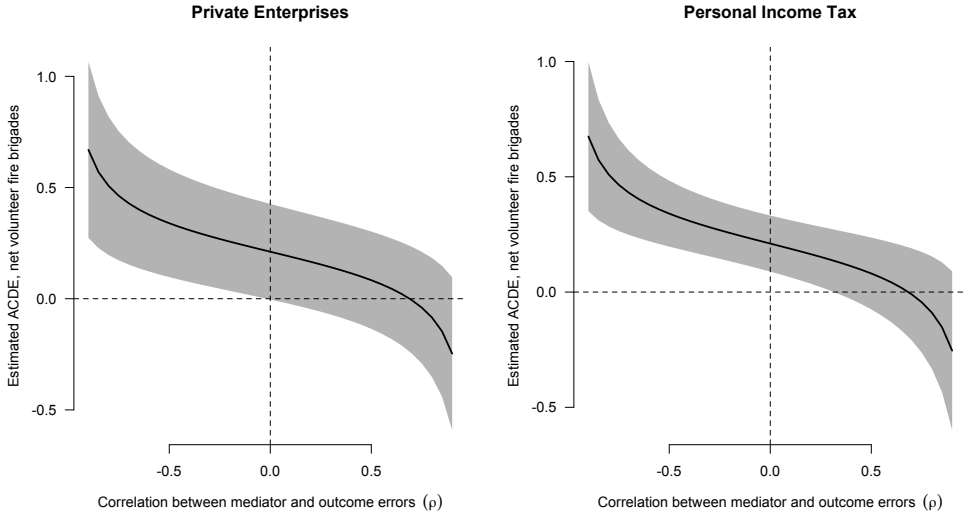


Table A16: G estimation first and second stage regressions for economic outcomes measured in 1998. Odd-numbered models are first-stage regressions with selection and post-treatment bias. Even-numbered models present the estimated direct effect of migrant heterogeneity after subtracting the proxies for voluntary public goods provision or fiscal capacity.

	ln(Enterprises per 1000 people)				ln(Personal Income Tax)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Volunteer Fire Brigades	−0.01*				−0.002			
	(0.002)				(0.002)			
log(Property Tax Revenues)			0.13**				0.02+	
			(0.02)				(0.01)	
Migrant Diversity	0.27**	0.21+	0.20**	0.15	0.28**	0.23**	0.27**	0.22**
	(0.08)	(0.11)	(0.08)	(0.10)	(0.07)	(0.06)	(0.07)	(0.06)
Intermediate Covariates	✓		✓		✓		✓	
✓		✓						
Pre-treatment Covariates	✓	✓	✓	✓	✓	✓	✓	✓
✓	✓	✓	✓					
District FE	✓	✓	✓	✓	✓	✓	✓	✓
✓	✓	✓	✓					
Moran Eigenvectors	✓	✓	✓	✓	✓	✓	✓	✓
✓	✓	✓	✓					
N	599	599	595	595	599	599	595	595
Adjusted R^2	0.76	0.66	0.78	0.66	0.63	0.56	0.63	0.56

+ p<0.1; * p<0.05; ** p<0.01

D.5 Instrumental variable analysis

I argue that reliance on formal versus informal institutions is the key mechanism through which historical heterogeneity affects contemporary economic outcomes. If we believe that heterogeneity produced by the population transfers is truly exogenous and does not affect economic outcomes through channels other than local institutional variation, then we can use *Migrant Diversity* as an instrument for the effect of institutional differences on economic outcomes. This section discusses additional assumptions that this approach requires and then implements the 2SLS analysis.

The key assumptions for instrumental variables estimation include exogeneity, excludability, and nonzero first stage.[7] The paper argues that conditional on covariates *Migrant Diversity* is a product of arbitrary decisions by local officials and is exogenous to local socio-economic conditions. Satisfying the exogeneity assumption is sufficient for the causal interpretation of the reduced form regressions of economic outcomes on *Migrant Diversity* and covariates, presented in the main body of the paper. The first-stage is also nonzero because *Migrant Diversity* predicts reliance on formal versus informal institutions for the provision of public goods, as shown in the paper. More problematic is the excludability assumption. The exclusion restriction requires that the cultural composition of resettled communities affect private economic activity only through the first-stage channel (i.e., by changing their reliance on formal versus informal institutions), conditional on the historical covariates that influenced the resettlement patterns. Excludability is not implausible because more than fourty years have passed since the population transfers. Today, the majority of communities formed by diverse migrant groups in the 1940s are well integrated and self-identify as Polish. Furthermore, much of the literature has found negative effects of heterogeneity on economic outcomes, which suggests that the violation of the exclusion restriction whereby migrant diversity would have a direct effect on economic outcomes would bias results toward zero. At the same time, it is possible that the mixing of cultures gave rise to a more individualistic or pro-business culture or affected selection into specific occupations. Relatedly, skill complementarities may have outlasted the cultural differences between migrants from different regions, producing a direct positive effect on productivity. Both of these possibilities cannot

[7] Other relevant assumptions are motonicity (those affected by heterogeneity are affected in the same way) and non-interference.

be ruled out and may violate the exclusion restriction.

Instrumental variable analysis presented below should be interpreted with these important caveats in mind. In contrast to the paper, where *Migrant Diversity* is the main causal variable, the instrumental variable analysis focuses on the reliance on informal institutions, proxied as the density of volunteer fire brigades, as the main cause.[8] Model 1 in Table A17 presents first-stage regression estimates. Models 2-3 present results of the OLS regression of entrepreneurship and personal income tax per capita on *Volunteer Fire Brigades* and historical covariates. The coefficient on *Volunteer Fire Brigades* is negative and statistically significant. Models 4-5 present second-stage regression results. Instrumenting for reliance on informal institutions for public goods provision with *Migrant Diversity* yields large and statistically significant estimates of the impact of institutional differences on incomes and entrepreneurship rates. The 2SLS estimates are larger than the OLS estimates, which may be because they capture local average treatment effects or because there is measurement error (e.g., not all volunteer fire brigades registered). The results are similar when the outcome variables are measured in alternative years. 2SLS regression analysis thus supports the argument about the relationship between reliance on informal institutions and economic outcomes. However, its interpretation hinges on the validity of exclusion restriction.

Table A17: Instrumental Variable Analysis of Migrant Diversity and Economic Outcomes Measured in 2000.

	Volunteer Fire Brigades	ln(Income)	ln(Enterprises)	ln(Income)	ln(Enterprises)
	First Stage	OLS		2SLS	
	(1)	(2)	(3)	(4)	(5)
Migrant Diversity	-4.03^{**}				
	(1.48)				
Volunteer Fire Brigades		-0.01^{**}	-0.01^{*}	-0.07^{*}	-0.05^{+}
		(0.002)	(0.003)	(0.03)	(0.03)
Covariates	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓
N	611	611	611	611	611
Adjusted R^2	0.24	0.53	0.51	0.07	0.29

+ p<0.1; * p<0.05; ** p<0.01

[8]The data on municipal guards are available only for 2007 and later years and is argued to have resulted from differences in voluntary public goods provision.

E Business environment in 2005

An additional empirical implication of the argument is better governance in historically more heterogeneous communities. To test this hypothesis, I analyze data from the 2005 Business Environment and Enterprise Performance Survey (BEEPS) (2005), which captures variation in governance at a much more fine-grained level than other data (e.g, the Quality of Governance Institute's dataset). The survey was carried out in Poland and other post-Communist states by the European Bank for Reconstruction and Development and the World Bank in 1999, 2002, and 2005. The 2005 wave included questions about obstacles to doing business and contained precise geographic identifiers that were matched to data on migrant origins.

Interviewees were owners, directors, and managers of 975 companies engaged in services and manufacturing. Of these firms, 224 (23%) are located in the resettled territories and are used in the analysis. Approximately three quarters (73%) of the firms sampled were created following the passing of the 1988 Fundamental Law on Economic Activity, which sanctioned the establishment of private commercial firms and accorded them equal status with state-owned firms. Most sampled businesses (84%) were created "from scratch" rather than from the privatisation of state-owned enterprises. The majority of firms sampled are small and nearly all (96%) are private.

Respondents were presented with a list of potential obstacles to doing business and asked to estimate how "problematic" each obstacle was. The predominance of informal over formal enforcement and lower stocks of state capacity would potentially generate problems such as the inadequate "Functioning of the judiciary" or "Tax administration," "Corruption," and "Organized crime." I used these options, with values 1 (no obstacle), 2 (minor obstacle), 3 (moderate obstacle) and 4 (major obstacle) to create dummy variables, coded 1 if respondents identified a given problem as "major" or "moderate."

The 2005 BEEPs covers many other aspects of the business environment. However, only some questions pick up on meaningful variation in the Polish case. This is because the overall business environment is ranked highly, two years after the country's EU accession and 15 years after the start of the institutional reforms. For example, Q41 asks: "Thinking of unofficial payments/gifts that a

firm like yours would make in a given year, could you please tell me how often would they make payments/gifts for the following purposes ..." Respondents are then presented with 10 different scenarious in which bribery may be needed, from dealing with courts to getting electricity. Each response ranges from 1 (never) to 6 (always). However, the mean response by Polish firms for this question is 1.31 (sd=0.674), which ranks between "never" and "seldom."

The grouped nature of the data (several firms in specific locations) allows using multi-level regression analysis to evaluate the role of location-specific factors, such as *Migrant Diversity* and *Share Migrants* in 1948, while controlling for firm-level characteristics. I control for the following firm-level characteristics: sector (Service or Manufacturing), size ($<$ *50 Employees*, *50-249 Employees*, and $>$ *250 Employees*), and the origin of the firm (*Private from the time of start up; Private subsidiary of a formerly state-owned firm; Joint venture with foreign partner*). In addition to the historical group-level measures of heterogeneity, I control for the type of sampling unit (*Rural Municipality*, *City under 50,000* and *City with 50,000-250,000 people*) and for *Distance to the EU border* because the survey took place after the EU accession. I restrict the sample to privately-owned firms founded *after* the market transition.

Full regression results are presented in Table A18. Coefficients on *Migrant Diversity* are negative in all models and statistically significant at a 0.10 level for the functioning of the judiciary, organized crime, and a mean of all four obstacles. Figure A11 summarizes regression results graphically. The average marginal probability that a respondent identifies functioning of the judiciary or corruption as an obstacle falls by more than half when comparing homogeneous to most heterogeneous communities.

While far from definitive, since subjective assessments of obstacles to doing business could reflect perceptions of economic outcomes rather than actual institutional quality, the analysis suggests better governance in historically more heterogeneous communities.

At the same time, *Migrant Diversity* does not predict the reporting of tax rates, contract violations, or access to financing as problems to doing business (see Appendix Table A19). Some of these null findings may seem puzzling. Contract violations, in particular, may reflect the quality of the business environment. One reason for the statistically insignificant coefficient on *Migrant Diversity* in this

case may be that the rates of contract violations do not vary across heterogeneous or homogeneous communities. The theory predicts that contract violations would be enforced through different institutional channels in heterogeneous and homogeneous communities, not that heterogeneity would reduce the incidence of contract violations.

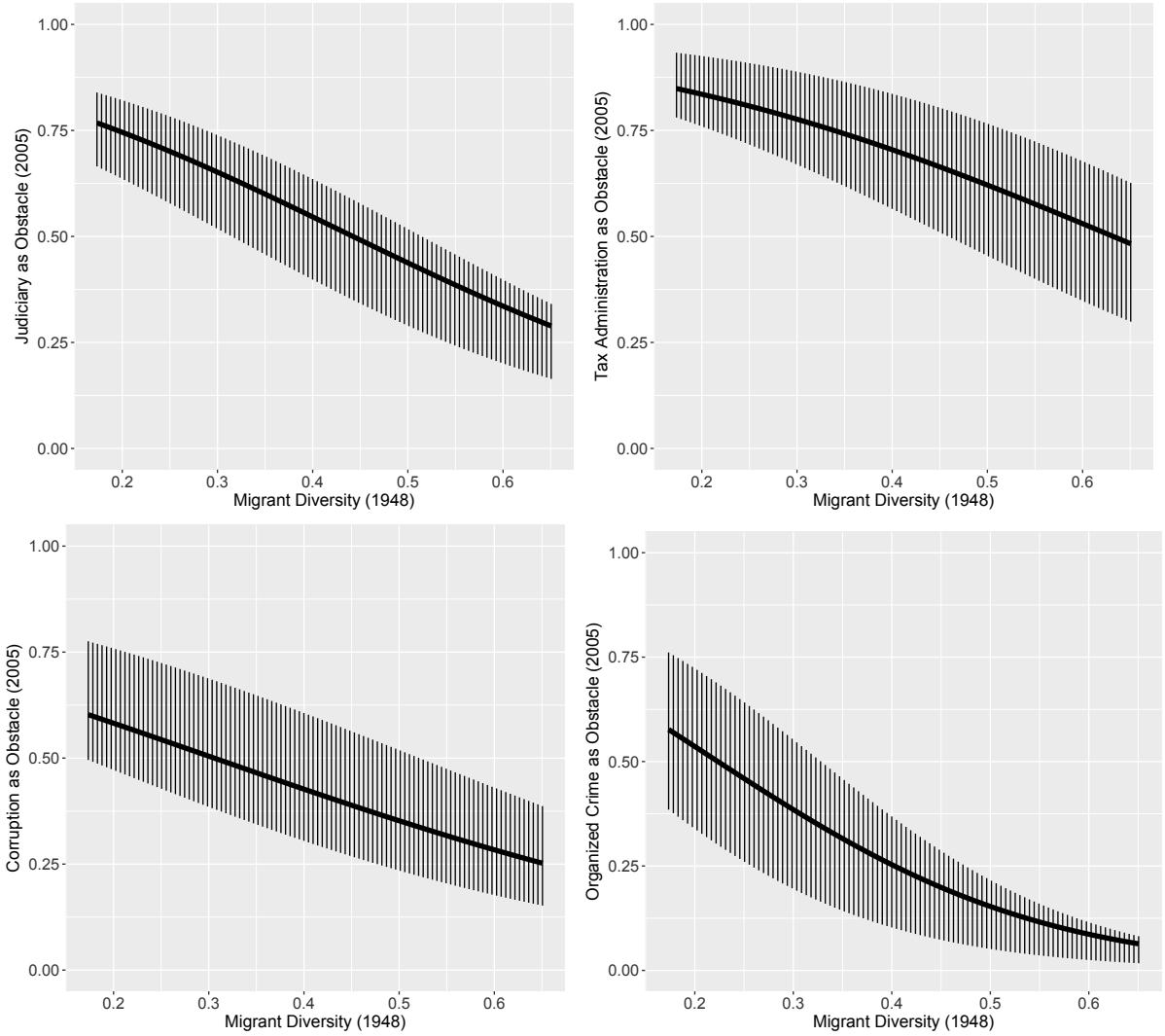


Figure A11: Average Marginal Probability of Identifying Specific Obstacles to Doing Business at Various Levels of Diversity. Probabilities are based on the multilevel logit models in Table A18.

Table A18: Migrant Diversity and Obstacles to Doing Business: (1) Functioning of the Judiciary; (2) Functioning of Tax Administration; (3) Corruption; (4) Organized Crime/Mafia; and (5) all four obstacles (mean). Multilevel Logit (1-4) and linear (5) models. Sample restricted to privately-owned firms.

	Judiciary	Tax Admin	Corruption	Org. Crime	All Four Obstables
	(1)	(2)	(3)	(4)	(5)
Migrant Diversity	-4.85^{+}	-4.45	-3.56	-7.49^{+}	-1.87^{+}
	(2.89)	(3.17)	(2.65)	(4.19)	(1.01)
Share Migrants	-0.10	-1.95	-0.22	-1.29	-0.49
	(1.19)	(1.23)	(1.12)	(1.53)	(0.40)
Size: >249 Employees	-0.56	1.39	1.14	1.86	0.22
	(1.28)	(1.45)	(1.21)	(1.40)	(0.45)
Size: <50 Employees	-0.01	-0.37	0.56	0.74	0.05
	(0.48)	(0.49)	(0.51)	(0.70)	(0.17)
Service Sector	-0.17	-0.17	-0.09	-0.04	0.06
	(0.40)	(0.38)	(0.41)	(0.55)	(0.15)
City: 50K-250K	0.47	0.50	1.22^{+}	0.54	0.20
	(0.75)	(0.63)	(0.66)	(1.05)	(0.26)
City: < 50K	0.09	1.40^{**}	-0.28	-0.99	0.09
	(0.58)	(0.51)	(0.46)	(0.89)	(0.20)
Founded: privatized	-0.79	-2.19^{+}	2.05	2.31	0.01
	(1.28)	(1.25)	(1.38)	(1.58)	(0.43)
Founded: from scratch	-0.25	0.59	1.34	0.40	0.15
	(0.98)	(0.92)	(1.18)	(1.28)	(0.34)
log(Distance to EU)	-0.60^{+}	-0.15	-0.56^{*}	-0.17	-0.14
	(0.32)	(0.25)	(0.26)	(0.50)	(0.11)
N	154	164	150	150	164
Log Likelihood	-99.12	-96.79	-90.48	-60.96	-203.15
AIC	222.24	217.57	204.95	145.92	432.31
BIC	258.68	254.77	241.08	182.05	472.61

+ p<0.1; * p<0.05; ** p<0.01

Table A19: Migrant Diversity and Obstacles to Business: (1) Tax rates; (2) Contract violations by customers or suppliers; (3) Access to Financing; or (4) Customs and Trade Regulations as a "Moderate Obstacle" or a "Major Obstacle" to Doing Business. Multilevel Logit. Sample restricted to privately-owned firms.

	Tax rates	Contract violations	Financing	Customs
	(1)	(2)	(3)	(4)
Migrant Diversity	0.54	4.08	1.32	−0.48
	(3.51)	(2.98)	(2.69)	(3.08)
Share Migrants	−0.97	1.16	−0.94	−0.58
	(1.47)	(1.09)	(1.08)	(1.21)
Size: $>$ 250 Employees	0.49	−0.71	1.51	−0.04
	(1.35)	(1.29)	(1.37)	(1.31)
Size: $<$ 50 Employees	0.04	0.23	0.36	−0.13
	(0.58)	(0.44)	(0.45)	(0.46)
Service Sector	−0.53	0.03	−0.60	0.40
	(0.50)	(0.37)	(0.40)	(0.40)
City: 50-250,000	−0.15	0.61	−1.80**	0.13
	(0.77)	(0.62)	(0.66)	(0.71)
City: <50,000	0.88	−0.36	−0.14	0.48
	(0.61)	(0.48)	(0.47)	(0.48)
Founded: privatized	0.28	−2.41+	0.62	−0.52
	(1.21)	(1.35)	(1.42)	(1.56)
Founded: from scratch	1.30	−0.83	2.69*	1.21
	(0.96)	(0.86)	(1.20)	(1.15)
log(Distance to EU)	−0.12	−0.05	−0.28	−0.15
	(0.32)	(0.24)	(0.26)	(0.25)
N	163	162	161	144
Log Likelihood	−69.59	−104.35	−93.07	−88.60
AIC	163.18	232.70	210.14	201.20
BIC	200.31	269.75	247.11	236.83

+ p<0.1; * p<0.05; ** p<0.01

F Alternative explanations

F.1 Cultural differences

Because the proportion of the largest cultural group at the municipality level is by construction correlated with *Migrant Diversity*, it could be that the presence of a specific cultural group, because of its human capital or attitudes to formal institutions, explains the findings. In particular, *Migrant Diversity* is negatively correlated with the share of the largest population group, migrants from Central Poland, and positively correlated with the shares of smaller population groups, migrants from the USSR (second largest group) and migrants from Western Europe (smallest group). If migrants from the USSR, for example, were more likely to rely on formal institutions or engage in entrepreneurship, their greater prevalence in heterogeneous areas could explain the results.

To address this concern, I regress the share of repatriates from the USSR, the share of migrants from Poland, and the share of migrants from Europe on the main social and economic outcomes in the paper. These variables are included in separate regressions because they are highly correlated. Table A20 shows that the coefficients on most group shares do not reach statistical significance. There are only two exceptions: the share of migrants from USSR is negative and statistically significant at a 10% level in Model 1, with *Volunteer Fire Brigades* as a DV, and the share of migrants from Europe is positive and statistically signifiant at a 5% level in Model 9, with Property Tax Revenue as a DV. Table A21 explores the relationship between the origin of migrants and economic outcomes in 1995. Again, the coefficients on the shares of each group do not reach statistical significance in most models. One exception is the negative relationship between the share of migrants from Central Poland and personal income tax per capita (Model 2). This analysis suggests that while cultural differences do matter for some social and economic outcomes, they do not provide a consistent explanation for all of the findings in the article.

As an additional test for the role of cultural heterogeneity above and beyond group characterictics, I use an alternative fractionalization index (*Cultural Diversity*), which combines all four population groups in formerly German territories (the indigenous population, migrants from Central Poland,

repartiates from the USSR, and migrants from Western Europe). To account for group differences, I include an indicator variable for the dominant group in each municipality, which includes four levels: *Largest: Central Poland, Largest: USSR, Largest: Europe, and Largest: Autochthonous*. Table A22 shows that the coefficients on *Cultural Diversity* are positive and statistically significant for all post-1989 economic outcomes even when the origins of the dominant group are accounted for.

Finally, Table A23 demonstrates that the results are broadly similar to results in the paper when using the alternative fractionalization index, composed of all four groups, instead of the *Migrant Diversity* index discussed in the article.

Table A20: Proportion of different migrant groups in 1948 and prevalence of organizations that rely on formal and informal enforcement in contemporary Poland.

	OLS Volunteer Fire Brigades			logistic Municipal Guard			OLS ln(Property Tax Revenue)		
	OLS			logistic			OLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Share from USSR	−1.90^{+}			1.12			−0.05		
	(1.06)			(0.80)			(0.16)		
Share from Centr. Poland		0.78			0.40			−0.07	
		(0.97)			(0.76)			(0.15)	
Share from Europe			0.47			2.17			0.93*
			(2.88)			(2.01)			(0.44)
Covariates	✓	✓	✓	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	611	611	611	611	611	611	607	607	607
Adjusted R^2	0.24	0.24	0.23				0.22	0.22	0.23
Log Likelihood				−243.09	−243.92	−243.50			
AIC				524.17	525.83	524.99			

+ p<0.1; * p<0.05; ** p<0.01

Table A21: Proportion of different migrant groups in 1948 and private economic activity in 1995.

	ln(Personal Income Tax per capita, 1995)			ln(Private Enterprises per 1000, 1995)		
	(1)	(2)	(3)	(4)	(5)	(6)
Share from USSR	0.02			0.03		
	(0.06)			(0.10)		
Share from Centr. Poland		-0.20^{**}			-0.09	
		(0.05)			(0.09)	
Share from Europe			0.05			0.10
			(0.15)			(0.27)
Covariates	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
N	611	611	611	611	611	611
Adjusted R^2	0.23	0.25	0.23	0.47	0.47	0.47

+ p<0.1; * p<0.05; ** p<0.01

Table A22: Diversity as Herfindahl (fractionalization) index with four groups and post-1989 Economic Outcomes. Alternative Specifications. OLS regression.

	ln(Personal Income Tax)			ln(Private Enterprises)		
	1995	1998	2000	1995	1998	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Cultural Diversity	0.12^{+}	0.20^{*}	0.18^{*}	0.30^{*}	0.32^{**}	0.31^{**}
	(0.07)	(0.08)	(0.08)	(0.13)	(0.10)	(0.10)
Largest: Indigenous	0.17	0.30^{*}	0.27^{*}	0.52^{**}	0.39^{*}	0.42^{**}
	(0.11)	(0.12)	(0.12)	(0.20)	(0.15)	(0.14)
Largest: Central Poland	0.04	0.18	0.21^{+}	0.40^{*}	0.31^{*}	0.37^{**}
	(0.10)	(0.11)	(0.12)	(0.18)	(0.14)	(0.14)
Largest: USSR	0.05	0.18	0.23^{*}	0.39^{*}	0.29^{*}	0.33^{*}
	(0.10)	(0.11)	(0.12)	(0.18)	(0.14)	(0.14)
Covariates	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
N	611	611	611	611	611	611
Adjusted R^2	0.24	0.35	0.52	0.47	0.52	0.52

+ p<0.1; * p<0.05; ** p<0.01

Table A23: Diversity as Herfindahl Index with four groups and post-1989 Economic Outcomes. Alternative Specifications. OLS regression.

	ln(Personal Income Tax per capita)			ln(Private Enterprises per 1000)		
	1995	1998	2000	1995	1998	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Cultural Diversity	0.08	0.14^{+}	0.17*	0.22^{+}	0.26**	0.25**
	(0.07)	(0.08)	(0.08)	(0.12)	(0.10)	(0.09)
Covariates	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
N	611	611	611	611	611	611
Adjusted R^2	0.23	0.35	0.52	0.47	0.51	0.51

+ p<0.1; * p<0.05; ** p<0.01

F.2 Human capital

Table A24 examines whether the findings can be explained by higher human capital in more heterogeneous communities. The dependent variable in Models 1-3 is *Share of the Population with Higher Education* according to the 1978, 1988, and 2002 censuses. In Model 4, the dependent variable is *Educational Diversity*, Herfindahl index computed from the shares of the population with different levels of education (higher, sectonary, vocational, and primary). Models 5-8 regress the logarithm of *Personal Income Tax* per capita and logarithm of *Private enterprises* per 1,000 people on the post-treatment proxies for the level of education and for educational diversity, *Migrant Diversity* and *Share Migrants*, and the remaining covariates.

Table A24: Migrant Diversity and Human Capital. OLS Regression.

	Share w/ Higher Edu			Edu Diversity	ln(Pers. Income Tax)		ln(Private Enterprises)	
	1978	1988	2002	1988	1995	1995	1995	1995
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Migrant Diversity (1948)	−0.003	−0.01	0.02*	0.65	0.21**	0.21**	0.33**	0.28*
	(0.01)	(0.03)	(0.01)	(0.46)	(0.08)	(0.08)	(0.13)	(0.14)
Share Migrants (1948)	0.002	0.003	0.004	2.31**	−0.20**	−0.20**	−0.06	−0.11
	(0.004)	(0.02)	(0.01)	(0.33)	(0.06)	(0.06)	(0.09)	(0.10)
Share w/ Higher Edu (1988)					0.02		2.45**	
					(0.13)		(0.21)	
Edu Diversity (1988)						−0.002		0.03*
						(0.01)		(0.01)
Covariates	✓	✓	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓	✓	✓
N	576	611	610	611	611	611	611	611
Adjusted R^2	0.48	0.54	0.58	0.45	0.25	0.25	0.57	0.47

+ p<0.1; * p<0.05; ** p<0.01

F.3 Occupational heterogeneity

Data on occupational differences between migrants from different regions are unavailable. However, the 1950 census lists population in various occupation categories at the county level. Occupations included in the census are agriculture, forestry, industry, construction, administration and banking, communications, trade, public services, health, sciences, and other. I use the proportion of population in each category data to calculate occupational herfindahl index at the county level. Mean levels of

Migrant Diversity at the county level are indeed correlated with *Occupational Heterogeneity* in 1950, but the correlation is weak (ρ=0.20) and *Migrant Diversity* explains only 0.03 of the variance in *Occupational Heterogeneity*. Table A25 regresses occupations on *Migrant Diversity* and covariates. We see that the statistical association between the two variables disappears once we control for *Share in Industry in 1939* or *Distance to the Railway*, which are included as covariates in the main analysis because they are pre-treatment. More important, Table A26 shows that the original results do not change when the post-treatment county-level occupational herfindahl index is included in regression models. Thus, greater occupational heterogeneity in more diverse communities does not explain the main findings in the article.

Table A25: Diversity in 1948 and Occupational Heterogeneity at the County Level in 1950.

	Occupational Heterogeneity (1950)			
	(1)	(2)	(3)	(4)
Migrant Diversity	0.21*	0.07	0.06	0.05
	(0.09)	(0.09)	(0.09)	(0.09)
Share Resettled	−0.11*	−0.11*	−0.08+	−0.13**
	(0.04)	(0.04)	(0.04)	(0.04)
Share in Industry			0.47**	0.10
			(0.10)	(0.12)
Distance to Railway		−0.01**		−0.004*
		(0.002)		(0.002)
Share Urban				0.20**
				(0.04)
log(Population)				0.001
				(0.02)
Distance to Border				−0.0001
				(0.0002)
N	123	123	121	121
Adjusted R^2	0.07	0.15	0.21	0.41

+ p<0.1; * p<0.05; ** p<0.01

Table A26: OLS Regression of Post-1989 Economic Outcomes on Posttreatment *Occupational Heterogeneity* Measure.

	ln(Personal Income Tax)			ln(Private Enterprises)		
	1995	1998	2000	1995	1998	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant Diversity	0.19*	0.24**	0.28**	0.31*	0.25*	0.24*
	(0.08)	(0.08)	(0.09)	(0.14)	(0.11)	(0.11)
Share Migrants	−0.24**	−0.08	0.06	−0.03	0.10	0.13
	(0.06)	(0.06)	(0.07)	(0.10)	(0.08)	(0.08)
Occupational Heterogeneity	−0.23*	0.14	0.13	0.16	0.30*	0.21
	(0.10)	(0.11)	(0.11)	(0.18)	(0.14)	(0.14)
Covariates	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
N	603	603	603	603	603	603
Adjusted R^2	0.26	0.32	0.50	0.46	0.51	0.51

+ p<0.1; * p<0.05; ** p<0.01

F.4 State policy

Did more heterogeneous communities receive preferential treatment from the state during the communist period? Results in Table A27 indicate that levels of state-provided public goods did not vary systematically across levels of heterogeneity. Furthermore, municipalities with more and less heterogeneous populations did not differ in levels of employment in socialized agriculture or industry and had similar budgets (i.e. heterogeneity did not attract additional state subsidies).

Table A27: *Migrant Diversity* and State Policy in 1980-1982. Dependent variables are (1) Libraries per 1,000 people; (2) Schools per 1,000 people; (3) Employed in Collectivized Agriculture per 1,000 people; (4) Employed in Nationalized Industry per 1,000 people; and (5) Municipal earnings per capita (in Złoty); (6) Compensatory subsidies. OLS Regression.

	Libraries	Schools	Agriculture	Industry	Municipal Earnings	Subsidies
	(1)	(2)	(3)	(4)	(5)	(6)
Migrant Diversity	−0.002	−0.04	−1.31	14.61	−0.31	−0.14
	(0.06)	(0.08)	(14.40)	(31.94)	(0.29)	(0.49)
Share Migrants	0.03	0.07	14.55	33.03	0.29	−0.59^{+}
	(0.05)	(0.06)	(10.44)	(23.11)	(0.21)	(0.35)
Socio-economic covariates	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
N	602	602	600	602	602	602
Adjusted R^2	0.39	0.38	0.66	0.32	0.14	0.22

+ p<0.1; * p<0.05; ** p<0.01

F.5 Sorting

Analysis in this section uses data on population mobility from the 1988 Census. The census distinguishes between the population living in a given municipality since birth, the population that arrived in the 1970s and the population that arrived in the 1980s. I use these data to proxy for sorting from one community into another. Next, I regress the sorting measures on *Migrant Diversity* and covariates. Results are presented in Table A28. Model 1 demonstrates a weakly negative bivariate relationship between *Migrant Diversity* and share of the population *Not living from birth*, though the coefficient does not reach statistical significance. However, the p-value increases further once the industrial profile of communities is accounted for (Model 2), and, with a full set of covariates, the relationship

between *Migrant Diversity* and *Not living from birth* becomes positive (Model 3). The results for *Arrived in the 1970s* and *Arrived in the 1980s* are similar (Models 4 and 5). This suggests that rather than sort into more homogeneous communities, which were initially better at providing local public goods, the population was sorting into the more heterogeneous communities, albeit at rather slow rates. The predicted differences in the share of population not living from birth at minimum versus maximum heterogeneity are just 5% of the population.[9]

Table A28: Migrant Diversity and Sorting. OLS regression.

	Share of the population				
	Not Living from Birth			Arrived (1970s)	Arrived (1980s)
	(1)	(2)	(3)	(4)	(5)
Migrant Diversity	−0.05	−0.02	0.06^{+}	0.04**	0.05**
	(0.03)	(0.03)	(0.03)	(0.01)	(0.02)
Share Migrants		0.17**	0.11**	−0.0003	−0.002
		(0.01)	(0.02)	(0.01)	(0.01)
Share in Industry		−0.11**	−0.07	−0.03*	−0.08**
		(0.03)	(0.04)	(0.01)	(0.02)
Other Covariates			✓	✓	✓
District FE			✓	✓	✓
N	627	609	609	609	609
Adjusted R^2	0.002	0.23	0.28	0.20	0.26

+ p<0.1; * p<0.05; ** p<0.01

[9]Note: all regressions control for the share of migrant population, which by definition is not living in a given municipality from birth.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects." *American Political Science Review* 110 (3): 512-529.

Banasiak, Stefan. 1965. "Settlement of the Polish Western Territories in 1945-1947." *Polish Western Affairs* 6 (1): 121-149.

Blusiewicz, Tomasz. 2015. "How the West Was Won: Pioneers, Settlers and Communism in the Polish 'Ziemie Odzyskane', 1945-1948." PIASA 73rd Annual Meeting.

Carnegie, Nicole Bohme, Masataka Harada, and Jennifer L. Hill. 2016. "Assessing Sensitivity to Unmeasured Confounding Using a Simulated Potential Confounder." *Journal of Research on Educational Effectiveness* 9 (3): 395-420.

Czerniakiewicz, Jan. 1987. *Repatriacja Ludnosci Polskiej z ZSRR 1944-1948*. Warszawa: Panstwowe Wydawnictwo Naukowe.

EBRD-World Bank. 2005. "Business Environment and Enterprise Performance Survey." http://go.worldbank.org/I4ZRB5O300.

Fong, Christian, Chad Hazlett, and Kosuke Imai. 2018. "Covariate Balancing Propensity Score for a Continuous Treatment: Application to the Efficacy of Political Advertisements." *Annals of Applied Statistics* 12 (1): 156-177.

Lach, Stanislaw. 1978. *Przemiany Spoleczno-Polityczne na Pomorzu Zachodnim 1945-1950*. Poznan: Wydawnictwo Poznanskie.

Rosenbaum, Paul R. 2002. *Observational studies*. New York, NY: Springer.

Stola, Dariusz. 2010. *raj bez wyjscia? Migracje z Polski 1949-1989*. Warszawa: Instytut Studiow Politycznych PAN.