ECONOMISTS FOR UKRAINE

WORKING PAPER SERIES

WORKING PAPER No. 33 JUNE 2026

LOCAL OFFICIALS IN IMPERIAL ADMINISTRATION AND LONG-RUN DEVELOPMENT IN UKRAINE

Marian Melnyk

I am grateful to Tymofii Brik, who advised this project in its early stages at the Kyiv School of Economics, and to Wim Vijverberg, whose continued guidance at the CUNY Graduate Center has been central to its development. I also thank Leticia Arroyo Abad for extensive feedback on an earlier draft, Natalya Naumenko for sharing Soviet-era electrification data, and Elena Besedina, Jonathan Conning, Timothy Goodspeed, Guillaume Haeringer, Sergii Kiiashko, Olga Kupets, Maksym Obrizan, and Leonid Peisakhin for helpful comments and suggestions, and participants in seminars at the CUNY Graduate Center, students at the Kyiv School of Economics, and participants at the 2025 Southern Economic Association conference for useful feedback. I gratefully acknowledge a stipend from the Review of Economic Studies.

ABSTRACT

Local Officials in Imperial Administration and Long-Run Development in Ukraine

Where more locals staffed courts, police, and administration, Ukrainian communities collect more tax and govern more effectively today. Regions entered the Russian Empire with different traditions of self-rule, leaving variation in who held local office. Using newly transcribed 1897 census records, I measure the share of local officials and link it to modern communities. A 10-percentage-point increase in local representation is associated with 7.1% higher local tax revenue per capita in 2021. An IV strategy supports the results, which survive a battery of robustness checks. I argue that self-governance built durable local institutions that persisted across imperial and Soviet rule.

JEL CLASSIFICATION: N13, O43, N43, P48

KEYWORDS: Persistence, institutions, governance, state capacity, development, empires, Russian Empire, Ukraine

Marian Melnyk
Department of Economics, CUNY
Graduate Center, New York
mmelnyk@gc.cuny.edu

Local Officials in Imperial Administration and Long-Run Development in Ukraine

Marian Melnyk*

June 2, 2026

Abstract

Where more locals staffed courts, police, and administration, Ukrainian communities collect more tax and govern more effectively today. Regions entered the Russian Empire with different traditions of self-rule, leaving variation in who held local office. Using newly transcribed 1897 census records, I measure the share of local officials and link it to modern communities. A 10-percentage-point increase in local representation is associated with 7.1% higher local tax revenue per capita in 2021. An IV strategy supports the results, which survive a battery of robustness checks. I argue that self-governance built durable local institutions that persisted across imperial and Soviet rule.

JEL codes: N13, O43, N43, P48

*PhD Candidate, Department of Economics, The Graduate Center, City University of New York. Email: mmelnyk@gc.cuny.edu.

I am grateful to Tymofii Brik, who advised this project in its early stages at the Kyiv School of Economics, and to Wim Vijverberg, whose continued guidance at the CUNY Graduate Center has been central to its development. I also thank Leticia Arroyo Abad for extensive feedback on an earlier draft, Natalya Naumenko for sharing Soviet-era electrification data, and Elena Besedina, Jonathan Conning, Timothy Goodspeed, Guillaume Haeringer, Sergii Kiiashko, Olga Kupets, Maksym Obrizan, and Leonid Peisakhin for helpful comments and suggestions, and participants in seminars at the CUNY Graduate Center, students at the Kyiv School of Economics, and participants at the 2025 Southern Economic Association conference for useful feedback. I gratefully acknowledge a stipend from the Review of Economic Studies. All remaining errors are my own. An Online Appendix follows the main text.

1 Introduction

Who staffs local government may matter long after formal institutions change. In 1897, some districts in the Russian Empire were primarily staffed by local officials in administration, courts, and police, while others were dominated by outside appointees. More than a century later, communities located in districts with greater historical local representation collect substantially more local tax revenue.

This paper asks whether historical experience of local self-governance within an empire explains contemporary economic development. Ukraine provides a sharp setting: its regions entered Russian imperial rule with different histories of autonomy, yet by the late eighteenth century all were governed under identical formal institutions. The variation that remained was informal but measurable: some districts retained locals in administrative, judicial, and police positions, others were staffed almost entirely by imperial appointees. This variation reflected not ethnic demography but institutional inheritance: the legacy of prior autonomous governance that had been formally abolished but not fully displaced.

I manually transcribe the 1897 Russian Empire census books, preserving the distinction between Ukrainian and Russian administrators that previous sources collapse into a single "Russian" linguistic category. The key explanatory variable is the share of local (Ukrainian) officials in district administration, courts, and police, measured continuously rather than as a binary classification of direct versus indirect rule. I link these historical data to 1,132 modern Ukrainian communities (hromadas, the amalgamated local-government units that gained genuine taxing authority under the 2014 decentralization reform) using a population-weighted spatial crosswalk. Ukraine's 2014 decentralization reform, which transferred genuine taxing authority to local governments, provides the institutional setting that makes this governance capacity fiscally legible. The reform did not activate a dormant legacy but revealed in fiscal data what centralized administration had kept invisible: lower mortality in 1929 shows the effect in the early Soviet period, and nighttime lights show it continuously from 1992 onward.

The identification argument rests on the historical origins of bureaucratic variation: the boundaries that determined which districts retained local administrators were shaped by seventeenth-century military and political events, not by local economic conditions, and subsequent imperial reforms were imposed uniformly from above. The primary design is OLS with province fixed effects, which absorbs all time-invariant differences across the eleven historical provinces and isolates within-province variation. A corroborative instrumental variable strategy uses log distance to Baturyn, the former capital of the Cossack Hetmanate, a semi-autonomous Ukrainian polity later absorbed into the Russian Empire. Baturyn was destroyed in 1709 and is today a settlement of about 2,400 people with no independent economic significance.

A 10-percentage-point increase in the historical share of local officials is associated with a 7.1% increase in per-capita local tax revenue in 2021. Sensitivity analysis using the Cinelli and Hazlett (2020) framework shows that an unobserved confounder would need to explain at least 12.1% of residual variation in both treatment and outcome to eliminate the baseline estimate. The result is not driven by any single province: leave-one-out estimates are stable across each of the ten main provinces, and Conley spatial standard errors leave the estimates significant. Ukrainian officials held real administrative authority: disaggregating by location, local representation in district capital cities (where the most influential positions are presumably concentrated) independently predicts contemporary outcomes in both OLS and IV designs. Corroborative IV estimates are roughly twice as large, consistent with attenuation in the baseline. The baseline IV strategy uses no fixed effects; with province fixed effects, the instrument's first stage substantially weakens: the F-statistic falls from 11.8 without them to 2.6 with them, paralleling the analogous first-stage collapse documented by Buggle and Nafziger (2021). This is one of the reasons I treat OLS as the primary design.

The effect extends beyond the primary outcome. Communities with historically higher shares of local officials exhibit lower transfer dependence, more active local tax administration, and higher discretionary spending, a broader pattern of effective fiscal governance. The

mechanism appears to operate through governance institutions and a more egalitarian social structure: districts with higher historical local representation had significantly lower land inequality in 1905, fewer elite domestic servants in 1897, and lower crude death rates in 1929. Human and physical capital channels are ruled out: literacy, industrialization, Soviet electrification, and road density are all unrelated to historical self-governance. The Soviet state itself treated these governance traditions as consequential: districts with higher local representation were collectivized significantly more intensively by late 1932, a revealed-preference argument that these institutions were still visible three decades after the census.

This paper documents persistence of historical local self-governance under empire, measured by the share of local officials in district administration, courts, and police in 1897. Most closely related is Iyer (2010), who shows that Indian districts under indirect British rule outperform those under direct administration; here the variation is continuous rather than binary, and operates within formally uniform imperial institutions rather than across distinct colonial regimes. Closest on mechanism is Dell et al. (2018), who show that Vietnamese areas long incorporated into the Dai Viet state sustain stronger local collective action and public goods today; my setting differs in that Ukrainian regions experienced different periods and degrees of prior autonomous governance, producing a continuous spatial gradient rather than a sharp frontier, which motivates a continuous treatment and proximity-based identification rather than a boundary discontinuity. Dell (2010) likewise identifies long-run effects of a historical colonial institution from within-country spatial variation. Bruhn and Gallego (2012) find political representation, not inequality or human capital, to be the most consistent intermediating channel between colonial activities and current development across the Americas. The contrast between local and imperial administration echoes the broader inclusive–extractive distinction (Acemoglu et al., 2002; Acemoglu & Robinson, 2012). Michalopoulos and Papaioannou (2013) document that pre-colonial ethnic institutions persist through uniform colonial and post-colonial administration in Africa. Becker et al. (2016) document persistent effects of the Habsburg Empire's more competent bureaucracy

on contemporary trust in courts and police across Eastern Europe, including parts of Ukraine straddling the former Habsburg border, and Grosfeld and Zhuravskaya (2023) similarly find that lands formerly under Habsburg rule remain more pro-democracy than those formerly under Russian rule. Within the Russian Empire specifically, Buggle and Nafziger (2021) and Markevich and Zhuravskaya (2018) document long-run consequences of coercive labor institutions (serfdom). This paper turns to a different feature of imperial institutional legacy: local administrative representation. I find that historical local self-governance appears to operate through governance quality, visible today in both stronger local public finances and higher local economic activity. Chaudhary and Iyer (2025) show that decentralization improves outcomes only when fiscal authority accompanies transferred responsibility; historical self-governance may be one source of the local capacity that makes such authority effective.

The paper proceeds as follows. Section 2 provides historical background, describes the data, and outlines the empirical strategy. Section 3 presents the main results and robustness checks. Section 4 documents persistence across the imperial, Soviet, and post-Soviet periods and examines mechanisms. Section 5 concludes.

2 Data and Empirical Strategy

2.1 Historical Background

Recent studies show that imperial legacies have lasting effects on trust, governance, and political behavior. Grosfeld and Zhuravskaya (2023) find that former Prussian and Austrian regions of Poland remain more pro-democracy than areas formerly under Russian rule. Similarly, Becker et al. (2016) demonstrate that former Habsburg territories exhibit higher trust in courts and police and lower perceived corruption. Peisakhin (2013) shows that people living in former Austrian settlements today hold more pro-European political attitudes than those on the former Russian side. These attitudinal differences complement the institutional evidence, suggesting that imperial legacies operate through multiple channels.

More broadly, Pop-Eleches (2015) demonstrates that decades of communist developmental efforts could not erase pre-communist intraregional differences, confirming that historical institutional legacies can persist through radical regime change. Collectively, these findings indicate that contrasting imperial institutions created enduring differences in governance quality, bureaucratic norms, and civic trust, even within present-day nation-states.

Ukraine provides an informative setting for examining how local self-governance persisted within the Russian Empire. Its regions entered imperial rule with starkly different experiences of autonomy. The Left Bank inherited the Cossack Hetmanate's institutions of elective offices and self-governance, established by Bohdan Khmelnytskyi's 1648 uprising and formally abolished in 1764.[1] The Right Bank lacked sustained autonomy: briefly part of the Hetmanate (1648–1667), it was returned to Polish-Lithuanian control by the Truce of Andrusovo and came under Russian rule only with the Second Partition of 1793. The South was administered directly by imperial authorities after conquest (Crimea 1783, Buh-Dnister territories 1792, Bessarabia 1812). By the late eighteenth century, all territories were under uniform imperial administration, yet by 1897 local administrative composition still mirrored this institutional heritage. A detailed chronology appears in Online Appendix A.

For more than a century before the 1897 census, no Ukrainian territory retained formal self-governance. Yet despite uniform imperial institutions, the composition of local bureaucracies varied considerably across regions. Communities above and below the median district-level share of local officials in 1897 are broadly similar on baseline characteristics: topography, soil quality, river density, and forest cover are balanced across the two groups (Table B.1 in Appendix B). Where they differ, it is in directions that cut against attributing the contemporary result to inherited prosperity: above-median communities are slightly less urbanized, slightly less literate, and have somewhat fewer merchants per 1,000 population. They are also located further north and east and are more ethnically Ukrainian.

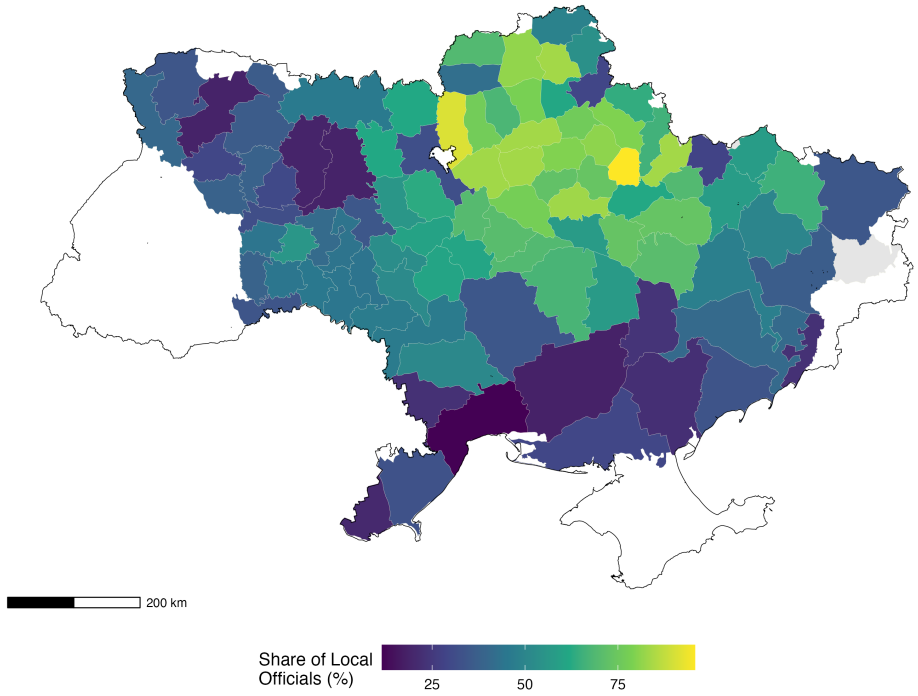

Figure 1: Share of local (Ukrainian) officials across 93 historical districts, 1897

The vast majority of Left Bank districts had a majority of Ukrainian officials in 1897; Right Bank and Southern districts generally had lower shares, though with considerable variation (Figure 1; boxplots by province in Online Appendix Figure OA.A.1). This variation is consistent with inherited governance experience: the boundaries of the Hetmanate and subsequent imperial reforms were imposed from above, not driven by local conditions, and local administrative composition continued to mirror historical traditions of self-governance.

Ethnicity and Self-Governance

A central feature of this paper's identification is the use of ethnic composition in local administration as a measure of self-governance. In the Russian Empire, ethnicity was not incidental to governance: it determined access to administrative positions. Imperial policies, including the Ems Decree (1876) and Valuev Circular (1863), restricted Ukrainian language use in public life and signaled a broader preference for Russian administrators in imperial governance. Although no systematic records document a formal replacement policy, the pattern of declining Ukrainian representation outside former Hetmanate territories is consistent with sustained centralization pressures. Where Ukrainians nevertheless retained positions in local administration, courts, and police by 1897, this likely reflects the depth of local institutional networks: where self-governance traditions had created generations of administrative experience, the imperial center found it more practical to work through existing local elites than to import Russian replacements for every district-level position. The ethnic composition of the bureaucracy was thus an institutional outcome, shaped by the historical geography of Cossack autonomy and the pace at which imperial centralization displaced local governance structures.

Three features support interpreting the share of local officials as a measure of self-governance rather than mere ethnic composition. First, the imperial ethnic hierarchy meant that Ukrainian representation in administration required active institutional channels: the same Cossack traditions that had sustained self-governance for over a century. In the absence

of such traditions, Ukrainian-majority districts on the Right Bank, where governance traditions were interrupted by the 1667 partition and subsequent Polish and then Russian direct rule, had far lower shares of Ukrainian officials, precisely because they lacked the institutional infrastructure to maintain representation in the bureaucracy. Second, the persistence of administrative ethnic composition long after formal abolition is consistent with theories of institutional transmission (Giuliano & Nunn, 2017): governance norms and administrative know-how, once embedded in local communities, can reproduce themselves through intergenerational transmission even without formal institutional support. Third, the empirical models control for the overall Ukrainian population share, so the explanatory variable captures governance involvement beyond what demographics alone would predict.

The instrumental variable approach below, exploiting proximity to Baturyn, reinforces this interpretation by isolating variation in Ukrainian officeholding driven by historical proximity to Cossack administrative institutions.

2.2 Data

To study the effect of historical self-governance on economic performance across the Ukrainian part of the Russian Empire, I construct a cross-sectional dataset at the community level (Amalgamated Territorial Communities, or hromadas) in 2021 (U-LEAD, 2025). The main dataset contains 1,132 observations with rich, granular information about Ukraine.

Outcome Variables

Local tax per capita, the primary dependent variable, captures the aggregate revenues from local taxes and fees collected by communities (see Online Appendix A for the Tax Code categories). The 2014 decentralization reform increased the share of revenues retained locally and incentivized communities to administer taxes effectively (Harus & Nivyevskyi, 2020; Romanova & Umland, 2019). Local tax revenues thus jointly reflect a community's economic activity and its governance capacity. Decentralization serves as a revelation mechanism: it

made legible in fiscal data a governance advantage that was continuously present but invisible under centralized administration.

I also employ alternative outcome variables to test potential mechanisms and provide robustness checks. Nighttime light intensity (LRCC-DVNL series, 1992–2020) from Tang et al. (2025) serves as an independent measure of local economic activity, allowing me to assess the persistence of self-governance effects across the entire post-Soviet period.

To capture the mechanism linking historical self-governance to contemporary outcomes, I follow the logic of Iyer (2010), who shows that indirectly governed regions in India had better governance quality. If districts in the Russian Empire with greater local representation developed more accountable administrations, this legacy should persist in the form of stronger fiscal governance today. Accordingly, I examine fiscal capacity measures that capture how communities raise revenue, manage transfer dependence, and allocate discretionary spending.

I complement local tax per capita with four measures of fiscal autonomy from the KSE Local Data Hub (Hatsko et al., 2025): the subsidy dependence rate, the transfer share, the local tax share excluding transfers, and per-capita discretionary spending on culture, physical education, and sport. These cover transfer dependence, the composition of own-source revenue, and discretionary spending power. Detailed definitions and the post-2014 fiscal framework are in Online Appendix A.

Explanatory Variables

The historical data are available at the district level, covering 93 *uezds* (districts) in the estimation sample, the third administrative tier within a province (*gubernia*). The novelty of my dataset is that it distinguishes observations by linguistic group. Other public sources and previous research aggregate Russians, Ukrainians, and Belarusians into a single linguistic category, "Russian," thereby erasing crucial ethnic differences. I manually transcribed the 1897 Russian Empire census books and preserved this distinction. The key explanatory

variable measures historical governance experience as the share of local (Ukrainian) officials employed in administrative, judicial, and police institutions in the 1890s (Trojnitskij, 1903—1905).

The main specification uses a continuous measure: the share of local officials among district administrators, court personnel, and police. This variable measures institutionalized local access to administrative office, not ethnic demography per se. As Figure 2 shows, office-holding was not mechanically proportional to local demography: Right Bank and Southern districts fall systematically below the 45-degree line despite Ukrainian demographic majorities. The governance-vs.-ethnicity distinction is tested directly in Section 3.

Control Variables

The models include contemporary controls (log community population, log community density), historical controls (urbanization, merchants per 1,000, literacy, industrialization, Ukrainian population share), and geographic controls (distance to Kyiv and the Black Sea coast, slope, river density, forest cover, chornozem share, and community centroid coordinates).[2] Standard errors are clustered at the historical district level. Table B.1 in Appendix B reports covariate balance; communities above and below the median share of local officials differ on several dimensions, motivating the control set. Table A.1 in Appendix A complements this by reporting the joint determinants of the treatment: geography alone explains 69% of cross-district variation, rising to 82% with the full historical control set.

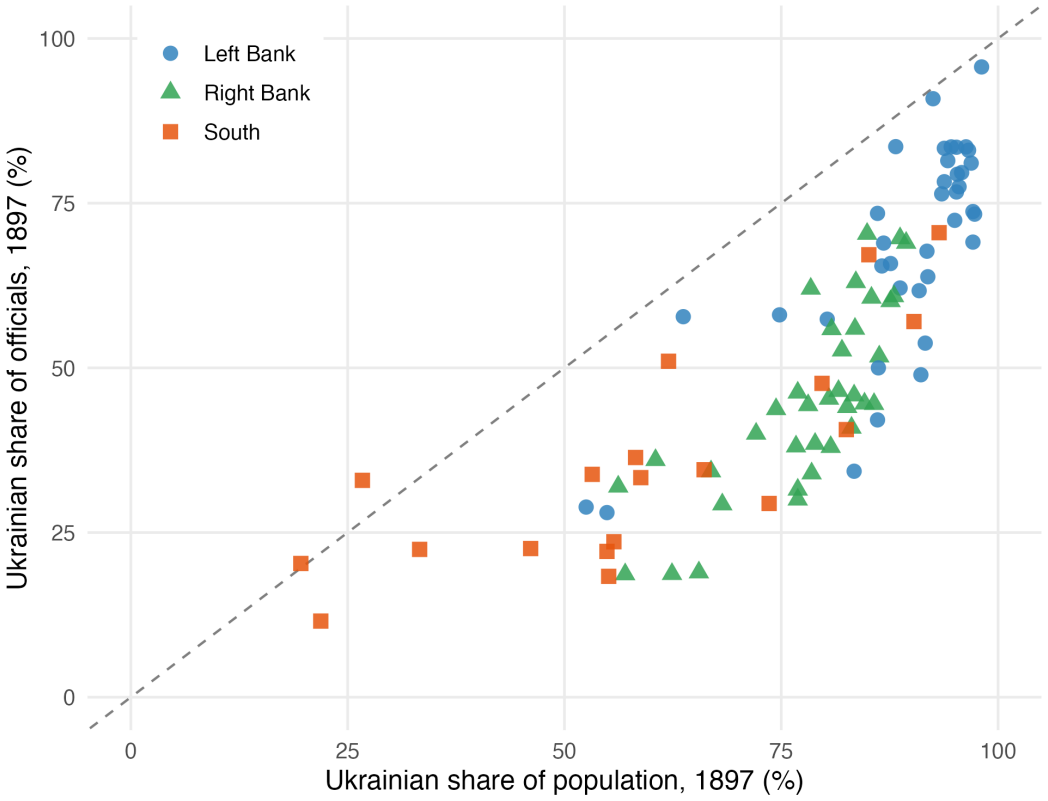

Figure 2: Ukrainian share of officials vs. population by district, 1897

Notes: Each point represents a historical district ($N = 93$). The dashed line marks proportional representation. Districts above the line had Ukrainians overrepresented in administration; districts below had Ukrainians underrepresented.

2.3 Matching and Descriptive Statistics

Table 1 reports summary statistics for the estimation sample. Local tax revenue averages 2,127 UAH per capita in 2021 (roughly 78 USD), and the share of local officials averages 52.5% across the 93 historical districts.

Table 1: Descriptive Statistics

Variable	N	Mean	SD	Min	Max
Community-level variables (ATCs)					
Local tax revenue p.c. (UAH, 2021)	1,132	2,127	1,572	349	20,733
Population (2020)	1,132	25,678	78,583	1,814	1,511,554
Population density	1,132	0.14	0.57	0.01	10.44
Log distance to Baturyn	1,132	12.72	0.57	8.16	13.51
District-level variables (1897 census)					
Share of local officials (%)	93	52.53	20.55	11.56	95.68
Ukrainian population share (%)	93	78.49	17.33	19.60	98.10
Literacy rate (%)	93	17.91	5.65	7.10	43.94
Urbanization rate (%)	93	11.84	5.56	5.98	28.86
Merchants per 1,000	93	2.29	1.90	0.11	10.55
Factories per 10,000 (1908)	93	1.38	1.62	0.02	7.92

Notes: The sample includes 1,132 communities (ATCs) matched to 93 historical districts of the Russian Empire. Community-level statistics are computed across ATCs; district-level statistics are computed across the 93 districts, each observed once. All communities within the same district share identical historical variable values. Local tax revenue includes the land tax, real estate tax, transport tax, unified tax, parking fee, and tourist fee. Distance to Baturyn is measured from community centroids in meters (log).

Online Appendix A reports extended descriptive statistics.

Because contemporary and historical borders differ, I match each community to the historical district containing more than 50% of its 2020 population, based on 2020 population density data (WorldPop & Bondarenko, 2020); Online Appendix A describes the matching procedure and exclusions in full. The estimation sample comprises 1,132 communities across 93 districts (12 communities per district on average).[3] All communities within the same district share identical historic values, so the effective historical variation lies across 93 districts; the community-level observations improve measurement of outcomes and controls.

Figure 3 plots the relationship between the historical share of local officials and the log of local tax revenues per capita in 2021. Each point represents a district-level mean, colored by macro-region, with the fitted line estimated on the underlying community-level data. Districts with a greater historical presence of local officials exhibit higher contemporary local tax revenues.

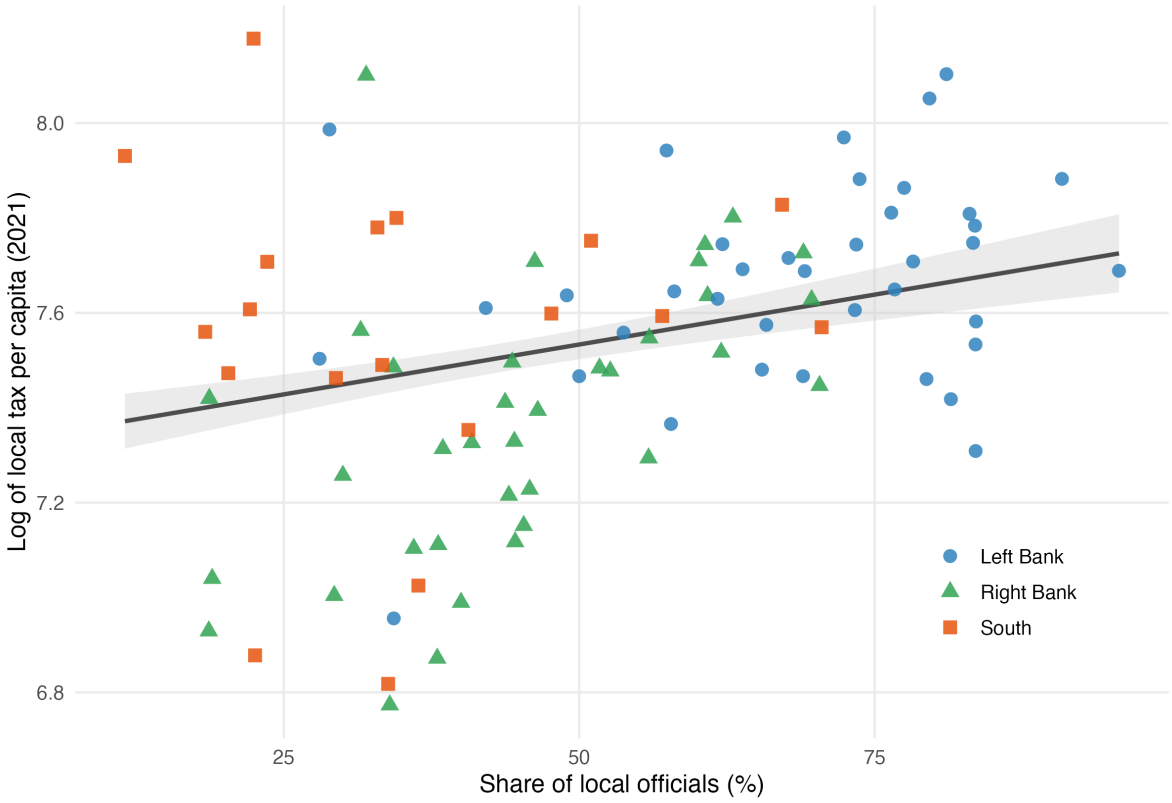

Figure 3: Historical share of local officials and contemporary local tax revenues per capita (log, 2021)

2.4 Empirical Strategy

The main specification uses the share of local officials in administration, courts, and police as a continuous measure of historical self-governance. As Section 2.1 discusses, the presence of Ukrainians in governance structures was determined largely by political and military events rather than by regional economic factors, providing a plausibly exogenous source of variation.

The empirical analysis begins with the following baseline specification:

$$\log(Y_i) = \alpha_0 + \alpha_1(\textit{Local Officials})_i + \alpha_2 X_i + \varepsilon_i \tag{1}$$

where Y_i denotes the per capita value of local tax revenues in community i or alternative outcome variables. The key independent variable, *Local Officials*, is the continuous share of local officials in district administration, courts, and police. The vector X_i includes contemporary, historical, and geographical covariates relevant to local economic performance, and ε_i is the error term. Standard errors are clustered at the historical district level (93 clusters), the level at which the treatment varies.

Instrumental Variable Strategy

A potential concern is that the share of local officials captures pre-existing advantages correlated with both self-governance and development, rather than the causal effect of governance experience. The areas that became centers of self-governance (like the Cossack Hetmanate) may have been systematically different before autonomy in geography, defense potential, or settlement patterns. If early Cossack communities formed in areas with more open land and better natural resources, those characteristics could influence both the emergence of self-rule and long-run economic outcomes. The primary way to mitigate the concerns is to control for geographic factors.

Another concern is that areas with stronger self-governance traditions may also have exhibited greater resistance or rebellion during the Soviet era, thereby prompting targeted

repression or unfavorable policy interventions that could have lasting adverse effects on local socioeconomic outcomes. Markevich et al. (2025) show that anti-Ukrainian bias in Soviet policies contributed to disproportionately high famine mortality among ethnic Ukrainians and that this bias was systematically applied across the Soviet Union. Because differential repression is itself a potential channel from the instrument to contemporary outcomes, I assess the exclusion restriction directly with the plausibly exogenous bounds of Conley et al. (2012) reported in Section 3, which allow the instrument a substantial direct effect on the outcome.

Thus, to mitigate potential endogeneity, I employ an instrumental variable (IV) approach. I use the distance to the town of Baturyn as an instrument for the share of local officials. Baturyn, historically the capital of the Cossack Hetmanate under Hetman Ivan Mazepa, functioned as a major center of Cossack self-governance until its destruction in 1709 and brief restoration before the final abolition of autonomy in 1764. Proximity to Baturyn thus captures historical exposure to institutions of local autonomy, under the assumption that areas closer to the historical Hetmanate institutional zone retained stronger traditions of local participation in governance. Baturyn is the natural anchor for this zone but is not the only historically motivated candidate. Today, Baturyn is a small town with a population of about 2,400, without economic significance. The identifying assumption is therefore that this proximity affects contemporary economic outcomes only through its historical influence on local administrative autonomy. Figure A.1 in Appendix A shows the location of Baturyn relative to the historical provinces. The instrument follows the logic of Becker and Woessmann (2009), who use distance to Wittenberg as an instrument for the spread of Protestantism. As Zhao (2023) notes, distance-to-a-point instruments can capture spatial gradients that are unrelated to the specific historical channel. Three features distinguish the present instrument from this concern. First, Baturyn was destroyed in 1709 and is today a town of about 2,400 with no economic significance, so it is not a cultural or economic center that could affect contemporary outcomes through non-governance channels. Second, replacing Baturyn with

alternative historically motivated anchors (the other two Hetmanate capitals, Chyhyryn and Hlukhiv, and the two main Sloboda Ukraine regimental centers, Sumy and Kharkiv) yields first stages and reduced-form coefficients in the same range as the Baturyn instrument; anchors at central places clearly outside the Hetmanate/Sloboda zone (Vinnytsia on the Right Bank, Rivne in Volhynia, Donetsk in eastern Ukraine, Mazyr in Belarus) instead produce weak first stages and statistically null reduced forms (Appendix Table B.4). That Kharkiv and Sumy are large modern cities while Baturyn is not reinforces the interpretation: the IV captures exposure to historical self-governance rather than proximity to any contemporary economic center. Third, the plausibly exogenous bounds of Conley et al. (2012), reported in Section 3, confirm that the IV estimate remains positive even when the instrument is allowed a substantial direct effect on the outcome.

Estimation equations:

$$\textit{Local Officials}_i = \pi_0 + \pi_1 \, \textit{Distance}_i + \pi_2 X_i + w_i \tag{2}$$

$$\log(Y_i) = \gamma_0 + \gamma_1 \, \widehat{\textit{Local Officials}}_i + \gamma_2 X_i + \eta_i \tag{3}$$

where *Distance* to Baturyn serves as the instrument in the first stage, and *Local Officials* is used as the explanatory variable in the second stage. X represents the control variables included in both stages.

The first-stage regression with the full set of controls yields an F-statistic of 11.8 ($p < 0.001$) for the instrument and 42.8 ($p < 0.001$) for the bare specification without controls. The bare F of 42.8 comfortably exceeds the conventional threshold; the preferred specification with controls yields F = 11.8, above the common rule-of-thumb of 10 (Staiger & Stock, 1997), indicating that the instrument passes the relevance test.[4] Adding fixed effects substantially weakens the first stage: F drops to 4.1 with macro-region fixed effects (three categories) and further to 2.6 with province fixed effects (eleven historical provinces), reflecting the limited cross-region variation exploited by the instrument; the preferred IV specification therefore

excludes fixed effects. The same smoothness rules out a boundary regression-discontinuity design (RDD), since the treatment does not jump at the historical boundary. Figure B.1 in Appendix B visualizes the first-stage relationship.

3 Results

3.1 Main Estimates

This section presents the estimation results, with standard errors clustered at the district level. Table 2 presents the main estimates linking historical governance experience to contemporary local economic outcomes. The key explanatory variable, Local Officials, is the continuous share of local officials in district administration, courts, and police. Columns (1)–(3) present OLS estimates with full controls and progressively finer fixed effects; columns (4)–(6) present IV estimates using distance to Baturyn as an instrument.

Column (1) includes the full set of geographic and historical controls without fixed effects. The coefficient on Local Officials is 0.0077, significant at the 1% level. Column (2) adds macro-region fixed effects (Left Bank, Right Bank, South); the estimate remains significant at 0.0056. Column (3), the preferred OLS specification, replaces region fixed effects with province fixed effects, absorbing all time-invariant differences across the eleven historical provinces. Relative to the no-fixed-effects baseline in column (1), the coefficient is essentially unchanged at 0.0071, confirming that the result is driven by within-province variation across districts, not by broad regional differences between the Left Bank and Right Bank. A 10-percentage-point increase in the historical share of local officials is associated with 7.1% higher contemporary local tax revenues per capita.

The OLS estimate with province fixed effects is the primary specification. Columns (4)–(5) present supplementary IV estimates using log distance to Baturyn as an instrument. The IV coefficients are consistently larger than OLS, consistent with classical measurement error in the manually transcribed census data and with omitted negative shocks from So-

viet repression that would bias OLS toward zero. Because the instrument operates primarily through cross-regional variation (first-stage F falls from 11.8 without fixed effects to 4.1 with macro-region fixed effects), the IV results should be interpreted as corroborative rather than definitive. This pattern is a structural feature of geographic instruments in historical settings: Buggle and Nafziger (2021), studying serfdom in the same Russian Empire districts, find that province fixed effects absorb most of the spatial variation in their treatment (joint geographic F drops from 21.9 to 2.3), and Markevich and Zhuravskaya (2018) cannot use their instrument once a collinear control is included. Thus, I follow the same approach: OLS with province fixed effects as the primary specification, IV as corroborative evidence that the cross-sectional OLS estimate (column 1, without fixed effects) is, if anything, attenuated. The IV estimates identify a local average treatment effect (LATE) for communities whose governance composition was most influenced by proximity to the Hetmanate's historical center; genuine treatment effect heterogeneity may also contribute to the larger IV magnitudes. Column (6) adds province fixed effects to the IV; the first-stage F falls to 2.6, below conventional thresholds, so this specification is reported for completeness only.

Table 2: OLS and IV Estimates

	OLS			IV		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Local Officials (%)	0.0077***	0.0056***	0.0071***	0.0165***	0.0199**	0.0309*
	(0.0016)	(0.0018)	(0.0024)	(0.0044)	(0.0088)	(0.0172)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
FE	—	Region	Province	—	Region	Province
Observations	1,132	1,132	1,132	1,132	1,132	1,132
Reduced form				-0.2142***	-0.1459***	-0.1995***
				(0.0699)	(0.0534)	(0.0668)
1st Stage F				11.8	4.1	2.6

Clustered (district) standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Dependent variable: log local tax revenue per capita (UAH, 2021). Sample: 1,132 amalgamated territorial communities (ATCs) matched to 93 historical Russian Empire districts. Local Officials (%) is the share of Ukrainian officials in district administration, courts, and police from the 1897 Russian imperial census. Columns (1)–(3) report OLS; columns (4)–(6) report IV using log distance to Baturyn (the historical Hetmanate capital) as an instrument. The *Reduced form* row reports the coefficient on log distance to Baturyn from a regression of log local tax revenue per capita on the instrument, with controls and fixed effects as in the corresponding IV column. Geographic controls include slope, river density, forest cover, distance to coast, distance to Kyiv, chornozem soil share, and community centroid coordinates. Other controls include log population, log population density, Ukrainian population share, urbanization (1897), merchants per 1,000, literacy, and factories per 10,000 (1908). Region FE: Left Bank, Right Bank, South. Province FE: 11 historical provinces (urbanization 1897 dropped due to collinearity).

Governance vs. Ethnicity

A natural concern is that the share of local officials proxies for ethnic composition rather than governance experience. All specifications control for the Ukrainian population share in 1897, but the two variables are correlated ($r = 0.80$). Table B.2 confirms this distinction: the Ukrainian share on its own is not a significant predictor of contemporary local tax revenue (column 2), and when entered jointly with governance the ethnicity coefficient turns slightly negative (column 3). Column 4 adds a centered interaction. The interaction coefficient is precisely estimated at zero, and the main effect is unchanged. The governance effect does not vary with ethnic composition, indicating that the main estimate is not driven by the correlation between the two variables.

3.2 Robustness

The main estimate is stable across a battery of checks reported in Appendix 5. Leave-one-province-out regressions show that the coefficient on Local Officials remains positive and statistically significant after sequentially excluding each of the ten historical provinces with multiple districts in the sample, with magnitudes close to the baseline (Table B.3).[5]

Three alternative inference procedures yield similar conclusions. Spatial standard errors of Conley (1999) at a 100 km cutoff leave the IV estimate unchanged and the OLS estimate significant at the 10% level (Online Appendix Table OA.C.1). Wild cluster bootstrap inference (Cameron et al., 2008) yields $p = 0.043$ with a 95% confidence interval of $[0.0003, 0.0130]$. The weak-instrument-robust Anderson–Rubin test rejects the null of zero IV effect at the 1% level (Online Appendix Table OA.C.2).

Two sensitivity exercises bound the role of unobservables. Following Cinelli and Hazlett (2020), an unobserved confounder would need to explain at least 12.1% of the residual variation in both treatment and outcome to eliminate the estimate; benchmarking against observed covariates indicates the result is reasonably resilient to omitted variable concerns. The plausibly-exogenous bounds of Conley et al. (2012) confirm that the IV estimate survives

moderate violations of the exclusion restriction.

Both capital-city and rural officials contribute to the result. Disaggregating the treatment by officials in district capital cities versus rural areas shows that both components independently predict contemporary outcomes, and the Baturyn instrument is, if anything, stronger for capital-city officials.

4 Persistence and Mechanisms

4.1 Persistence

The self-governance effect is visible across the imperial, Soviet, and post-Soviet periods. Already in the late imperial period, districts with more local officials had significantly lower land concentration in 1905 and fewer elite domestic servants, indicating a less hierarchical social structure consistent with Sokoloff and Engerman (2000)'s argument that egalitarian factor endowments promote inclusive institutions (Appendix Table C.1). The pattern extends into the early Soviet period: raion-level crude death rates in 1929 were significantly lower where local representation was higher, and urbanization in 1926 is marginally positive. District-level nighttime lights from 1992 through 2020 show the same positive effect: the coefficient on local-officials share is positive in every year and marginally significant ($p < 0.10$) in 24 of the 29 years, with a stable point estimate after early-1990s noise dissipates (Appendix Figure C.1). The pattern thus predates Ukraine's independence-era reforms: it is present at the start of the post-Soviet period, not produced by it.

The Soviet state itself treated these as institutionally distinct: districts with higher local representation were collectivized more intensively in 1932 despite identical procurement quotas and plan fulfillment rates (Appendix Table C.2). A 10-percentage-point increase in the historical share of local officials is associated with a 2.3-percentage-point higher collectivization rate, and famine mortality at the raion level tracked collectivization intensity closely (Markevich et al., 2025). Yet the governance traditions that provoked the Soviet response

ultimately proved more durable than the attempt to dismantle them. Ukrainian republican authorities, for their part, used the 1957 Sovnarkhoz reform to assert greater local control (Kibita, 2013), evidence that demand for local governance remained active four decades into Soviet rule. Taken together, this evidence indicates that local institutions survived Soviet homogenization rather than having been created by post-1991 reforms. Persistence here is observed at five intermediate periods: 1905, 1926, 1929, 1932, and 1992–2020; the transmission chain is empirically traceable.

4.2 Mechanisms

The main effect reflects a broader pattern of fiscal autonomy: districts with higher local representation also exhibit lower transfer dependence, more active local tax administration, and higher discretionary spending on culture, sport, and administration (Appendix Tables B.5, B.6). The channel appears to be institutional rather than physical, in line with Dell et al. (2018), who identify local collective action through civil society and local government as the central channel through which a historical state's legacy persists. Historical self-governance is unrelated to industrialization in 1908, Soviet electrification across five decades, or road density in the early 1990s (Online Appendix Table OA.B.1); human capital also shows no effect (Online Appendix Table OA.B.2). Table 3 summarizes the evidence.

Table 3: Summary of Mechanism Evidence

Finding	**Sign / Sig.**	**Interpretation**
Contemporary outcomes		
Fiscal capacity (4 measures)	3 of 4 sig.	Governance effect across multiple fiscal dimensions
Imperial-era structure		
Land Gini 1905	Negative (***)	Egalitarian land institutions
Elite servants 1897	Negative (**)	Less hierarchical social structure
Persistence through the Soviet period		
Urbanization 1926	Positive (*)	Adaptation to early Soviet modernization
Crude death rate 1929 (raion)	Negative (***)	Health advantage under early Soviet rule
Collectivization 1932 (raion)	Positive (**)	Soviet state treated regions as institutionally distinct
Procurement quotas	Null	Not wealth-based targeting
Fulfillment rates	Null	Not more intensive extraction
NTL 1992–2020 (district, yearly)	Positive (*)	Continuous post-Soviet persistence
NTL 2014–2020 (ATC, yearly)	Positive (***)	Confirmed at finer resolution
Channels ruled out		
Industrialization 1908	Null	No physical capital channel
Electrification 1928–1969	Null	No physical capital channel
Road density (1990s)	Null	No physical capital channel
Literacy, schools	Null	No human capital channel

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Taken together, the evidence is consistent. A 10-percentage-point increase in historical local representation is associated with 7.1% higher per-capita local tax revenue today; the relationship survives standard inference and sensitivity checks, the effect is visible across imperial, Soviet, and post-Soviet periods, and alternative channels (ethnicity, human capital, physical capital) do not account for it. What remains is a governance channel embedded in local administrative tradition.

5 Conclusion

This paper asks whether historical local self-governance under empire has persistent effects on contemporary economic development. Using novel data from the 1897 Russian Empire census, I measure self-governance as the share of local officials in district administration, courts, and police and link it to 1,132 Ukrainian communities in 2021. A 10-percentage-point increase in historical local representation is associated with 7.1% higher local tax revenue per capita. The effect reflects governance institutions, not human or physical capital: literacy, industrialization, electrification across five decades, and road density show no association with the treatment.

The transmission chain is empirically traceable: the self-governance signal appears at five intermediate periods (1905, 1926, 1929, 1932, and 1992–2020), not just at the two endpoints. These results indicate that the way local populations participated in governance under empire shaped the quality of local institutions that persist today, even after more than a century of uniform imperial administration and seven decades of Soviet homogenization.

The cross-sectional design limits dynamic analysis, and the instrumental-variable strategy identifies a local effect for districts in the historical core of Ukrainian self-governance. Yet the broader claim is general: local institutions can persist for generations through their administrators, outlasting successive central regimes.

Appendix A. Data

Table A.1: Determinants of Local Official Share

Dependent Variable:	Share of local officials (%)			
	Geography	+ Demographic	+ Historical	+ Region FE
Model:	(1)	(2)	(3)	(4)
Variables				
Log district population	-29.77***	-20.26***	-17.96***	-8.863
	(5.489)	(5.350)	(6.176)	(6.151)
River density	30.11	153.2*	163.6*	52.93
	(99.01)	(91.85)	(91.21)	(91.63)
Forest cover	41.64	24.34	7.287	21.75
	(26.58)	(22.26)	(25.88)	(24.26)
Terrain slope	0.0388	2.630	3.456	5.138
	(5.621)	(5.202)	(5.300)	(4.613)
Log dist. to coast	7.262	-0.4268	-3.936	-6.096
	(4.696)	(5.197)	(5.761)	(4.216)
Latitude	-3.614	-1.360	1.231	-0.1486
	(3.377)	(3.521)	(4.116)	(3.284)
Longitude	2.447***	1.133**	1.214**	0.0063
	(0.5352)	(0.4451)	(0.4876)	(0.5434)
Log dist. to Kyiv	-16.88***	-10.54***	-10.59***	-11.58***
	(2.947)	(2.121)	(2.090)	(1.979)
Chornozem share	20.43***	9.670	7.257	8.318
	(7.519)	(6.566)	(6.863)	(6.500)
Ukrainian share 1897 (%)		0.7124***	0.6198***	0.6175***
		(0.1043)	(0.1188)	(0.1047)
Urbanization 1897		65.29**	61.10**	
		(26.05)	(26.79)	
Literacy rate 1897			-61.02	-66.74*
			(46.64)	(39.78)
Merchant share			0.3573	0.0460
			(0.9391)	(0.7926)
Industrialization			-0.6349	-0.2672
			(0.7058)	(0.7471)
FE	—	—	—	Region
Observations	93	93	93	93
R^2	0.68611	0.81356	0.82303	0.84019

Heteroskedasticity-robust standard-errors in parentheses

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: OLS. Dependent variable: share of Ukrainian officials in district administration, courts, and police (1897, %). District-level sample, N=93. HC1 robust standard errors. Controls escalate from geography and soil alone (1) to macro-region fixed effects (4: Left Bank / Right Bank / South). *Urbanization 1897* is omitted in column (4) due to collinearity.

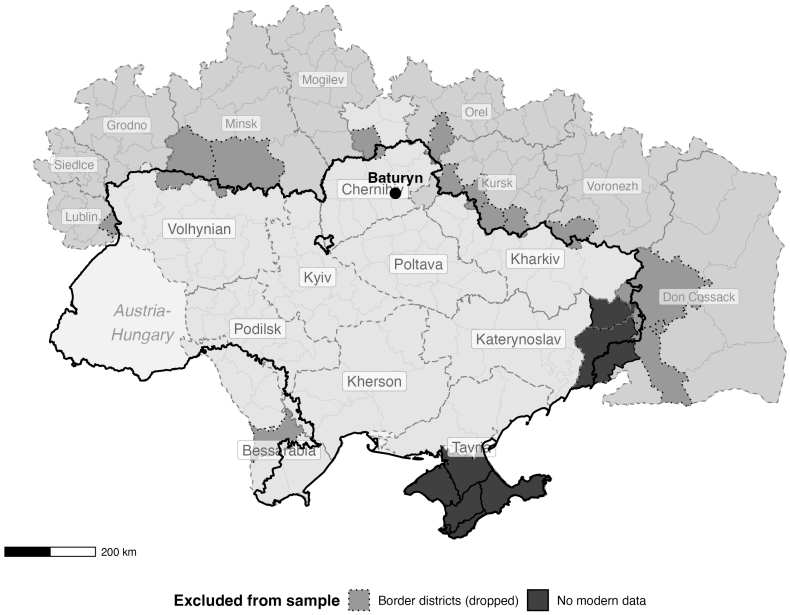

Figure A.1: Historical provinces, location of Baturyn, and excluded territories

Notes: Light grey areas show Ukrainian provinces at their full historical extent; darker grey shows neighboring non-Ukrainian provinces. The black dot marks Baturyn, the former Hetmanate capital used as the basis for the instrumental variable. Mid-grey fill with a dotted edge marks border districts dropped due to ambiguous overlap with modern Ukraine. Dark grey fill with a solid edge marks districts without modern ATC data (Crimea, Slovianoserbsk, Taganrog). The white area within Ukraine's borders (western regions) belonged to Austria-Hungary.

Appendix B. Robustness Tables

Table B.1: Covariate Balance by Historical Governance Intensity

	Below median		Above median		
	Mean	SD	Mean	SD	Norm. diff.
Population density	0.17	0.61	0.11	0.54	−0.10
Population	29,227	100,530	22,217	48,240	−0.09
Slope (mean)	0.59	0.38	0.65	0.36	0.18
River density	0.23	0.06	0.23	0.05	0.03
Forest cover	0.19	0.18	0.22	0.14	0.15
Log dist. to coast	4.94	1.50	5.58	0.50	0.57
Longitude	30.61	4.05	32.50	2.88	0.54
Latitude	48.74	1.75	49.47	1.12	0.50
Ukrainian pop. share 1897 (%)	58.19	16.25	85.63	7.97	2.14
Urbanization 1897 (%)	15.92	8.27	11.86	5.33	−0.58
Merchants per 1,000	3.49	2.86	2.17	1.25	−0.60
Literacy rate (%)	23.32	7.97	16.17	3.31	−1.17
Factories per 10,000 (1908)	1.91	1.95	1.81	2.01	−0.05
Dist. to Kyiv (log)	12.70	0.63	12.34	0.61	−0.59
Chornozem share	0.63	0.44	0.69	0.39	0.14
N	559		573		

Notes: Communities split at the median share of local officials (40.6%). Normalized differences computed as $(\bar{x}_1 - \bar{x}_0)/\sqrt{(s_1^2 + s_0^2)/2}$. Values above 0.25 indicate meaningful imbalance. All variables are pre-treatment or geographic covariates included in the main specification.

Table B.2: Historical Self-Governance vs. Ethnic Composition

	(1) Governance	(2) Ethnicity	(3) Both	(4) Interaction
Local Officials (%)	0.0055**		0.0071***	0.0072***
	(0.0022)		(0.0024)	(0.0024)
Ukrainian Pop. Share		-0.0012	-0.0048	-0.0049
		(0.0028)	(0.0031)	(0.0032)
Local Off. × Ukr. Pop.				-0.00000
				(0.00006)
Controls	Yes	Yes	Yes	Yes
FE	Province	Province	Province	Province
Observations	1,132	1,132	1,132	1,132
R^2	0.374	0.367	0.377	0.377

Clustered (district) standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Dependent variable is log of local tax revenue per capita in 2021. Column (1) includes the historical share of local officials without controlling for the Ukrainian population share. Column (2) replaces governance with the Ukrainian population share. Column (3) is the preferred main specification with both variables. Column (4) adds a centered interaction (both variables demeaned before interacting) to the specification in column (3); the Ukrainian population share is dropped from the control set to avoid collinearity with its centered counterpart in the interaction. All specifications include province fixed effects (11 historical gubernias) and the full set of demographic, geographic, and historical controls.

Table B.3: Effect of Local Officials on Local Taxes: Leave-One-Province-Out

Panel A. Provinces 1–5					
Model:	Podil (1)	Kyiv (2)	Volhyn. (3)	Kateryn. (4)	Poltava (5)
Local Officials (OLS)	0.0069***	0.0071***	0.0067**	0.0057**	0.0070***
	(0.0024)	(0.0025)	(0.0025)	(0.0025)	(0.0024)
Local Officials (IV)	0.0161***	0.0177***	0.0182***	0.0114***	0.0160***
	(0.0042)	(0.0055)	(0.0052)	(0.0036)	(0.0053)
Controls	Yes	Yes	Yes	Yes	Yes
OLS FE	Province	Province	Province	Province	Province
IV FE	—	—	—	—	—
Observations	1,033	997	940	990	1,028
Panel B. Provinces 6–10					
Model:	Kherson (6)	Kharkiv (7)	Tavria (8)	Chernih. (9)	Bessarab. (10)
Local Officials (OLS)	0.0058**	0.0069**	0.0074***	0.0113***	0.0053**
	(0.0024)	(0.0029)	(0.0024)	(0.0019)	(0.0020)
Local Officials (IV)	0.0155***	0.0220***	0.0166***	0.0190***	0.0228***
	(0.0046)	(0.0060)	(0.0046)	(0.0046)	(0.0076)
Controls	Yes	Yes	Yes	Yes	Yes
OLS FE	Province	Province	Province	Province	Province
IV FE	—	—	—	—	—
Observations	960	1,031	1,060	1,064	1,091

Clustered (district) standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Dependent variable: log local tax revenue per capita (UAH, 2021). Each column excludes all communities historically belonging to the named province from the estimation sample. Specification matches Table 2, columns (3) and (4): Local Officials (%) is the share of Ukrainian officials in 1897; controls are the full set listed in the notes to Table 2; OLS uses province fixed effects (remaining 10 provinces); IV uses log distance to Baturyn without fixed effects.

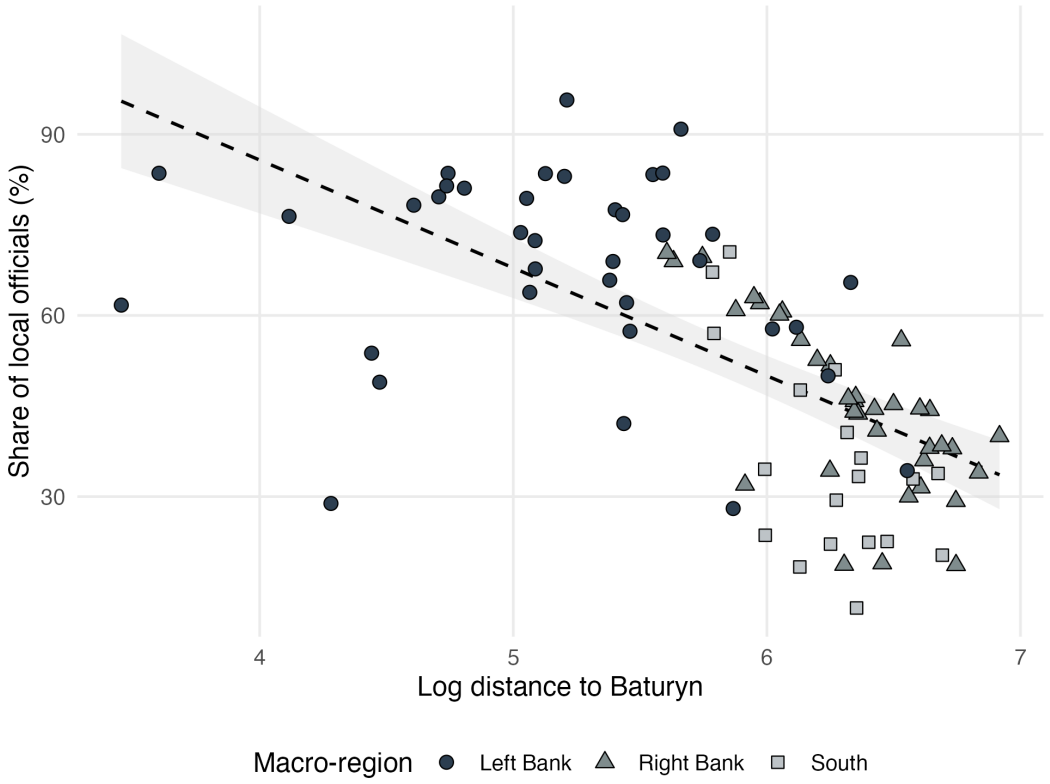

Figure B.1: First stage: Distance to Baturyn and share of local officials

Notes: Each point represents a historical district ($N = 93$). The dashed line shows the OLS fit across all districts; the shaded area is the 95% confidence interval. Point shapes and fill shades distinguish macro-regions (Left Bank, Right Bank, South).

Table B.4: Placebo IV: Alternative Spatial Anchors

	Coef. on *Local Officials*	Cluster SE (*povit*)	Conley SE (100 km)	F (cluster 100 km)	F (Conley)
Panel A. Cossack administrative centers					
Baturyn (1669-1708, 1750-1764)	0.0165***	(0.0044)	[0.0044]	11.84	9.01
Chyhyryn (1648-1669)	0.0149***	(0.0045)	[0.0043]	15.60	13.09
Hlukhiv (1708-1750)	0.0263***	(0.0082)	[0.0093]	5.93	3.48
Sumy (Sloboda center)	0.0230***	(0.0053)	[0.0051]	6.11	5.09
Kharkiv (Sloboda center)	0.0208***	(0.0052)	[0.0044]	9.56	7.59
Panel B. Placebo anchors outside the Cossack zone					
Vinnytsia (Right-Bank)	0.0115*	(0.0059)	[0.0056]	3.74	3.53
Rivne (Volhynia)	0.0096	(0.0087)	[0.0067]	2.31	2.32
Donetsk (industrial Donbas)	0.0544*	(0.0292)	[0.0270]	2.74	2.57
Mazyr (Belarus)	0.0116	(0.0110)	[0.0128]	3.54	3.63
Observations			1,132		
Controls			Yes (15 covariates)		
FE			None		

Clustered (povit) standard-errors in parentheses; Conley (100 km) in brackets
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Each row replaces the instrument in Table 2, column (4), with the log distance from each ATC centroid to the anchor point listed. Specification, sample, and controls match column (4) exactly: full vector of geographic, soil, and 1897-historical controls; no fixed effects. The IV coefficient is identical under either inference choice; only the standard errors and first-stage Wald F differ. Panel A reports anchors at the historical Cossack administrative centers: the three Hetmanate capitals (Chyhyryn, 1648–1669; Baturyn, 1669–1708 and 1750–1764; Hlukhiv, 1708–1750) and the two main Sloboda Ukraine regimental centers (Sumy, Kharkiv). Panel B reports placebo anchors at central places outside the Cossack zone: Vinnytsia (Right-Bank, Polish-Lithuanian), Rivne (Volhynia, Polish-Lithuanian), Donetsk (industrial Donbas, settled post-1869), and Mazyr (Belarus).

Table B.5: Historical Self-Governance and Fiscal Capacity

Panel A. OLS				
	Subsidy Dep. Rate (1)	Transfer Share (2)	Local Tax Share (3)	log Own Rev. p.c. (4)
Local Officials (%)	-0.0010^{***}	-0.0015^{***}	0.0005	0.0064^{***}
	(0.0003)	(0.0005)	(0.0005)	(0.0016)
Controls	Yes	Yes	Yes	Yes
FE	Province	Province	Province	Province
Observations	1,132	1,132	1,132	1,132
R^2	0.286	0.375	0.262	0.302
Panel B. IV				
	Subsidy Dep. Rate (5)	Transfer Share (6)	Local Tax Share (7)	log Own Rev. p.c. (8)
Local Officials (%)	−0.0008	-0.0028^{***}	0.0033^{***}	0.0032
	(0.0006)	(0.0010)	(0.0009)	(0.0033)
Controls	Yes	Yes	Yes	Yes
FE	—	—	—	—
Observations	1,132	1,132	1,132	1,132
R^2	0.260	0.354	0.253	0.261
1st Stage F	11.8	11.8	11.8	11.8

Clustered (district) standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Subsidy Dep. Rate = equalization grants as a share of total revenue; Transfer Share = all intergovernmental transfers as a share of total revenue; Local Tax Share = local taxes and fees (property, excise, unified, parking, tourist, eco) as a share of own-source revenue; log Own Rev. p.c. = log own-source revenue per capita (UAH). Source: KSE Local Data Hub (Hatsko et al., 2025), 2021 budget data. All models include demographic, geographic, and historical controls. OLS includes province fixed effects (11 historical gubernias; urbanization 1897 dropped due to collinearity); IV uses log distance to Baturyn as instrument without fixed effects.

Table B.6: Historical Self-Governance and Contemporary Governance Outcomes

	log Culture & Sport p.c. (1)	log Capital Exp. p.c. (2)	log Admin p.c. (3)
Local Officials (%)	0.0050***	0.0045	0.0025**
	(0.0019)	(0.0039)	(0.0012)
Controls	Yes	Yes	Yes
FE	Province	Province	Province
Observations	1,132	1,132	1,132
R^2	0.183	0.153	0.362

Clustered (district) standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: DVs in logs. Culture & Sport = per capita spending on culture, physical education, and sport (budget functional code 08). Capital Exp. = per capita capital expenditures (economic codes 31–32). Admin = per capita spending on administrative apparatus (functional code 01). Source: KSE Local Data Hub (Hatsko et al., 2025), 2021 expenditure data. All models include demographic, geographic, and historical controls with province fixed effects (11 historical gubernias; urbanization 1897 dropped due to collinearity).

Appendix C. Persistence and Mechanism Evidence

Table C.1: Historical Self-Governance and Intermediate Outcomes

	Land Gini 1905	Elite servants per 1k	Urban 1926	CDR 1929 (raion)
	(1)	(2)	(3)	(4)
Local Officials (%)	-0.002***	-0.003**	0.003*	-0.052***
	(0.001)	(0.001)	(0.001)	(0.018)
Observations	89	93	81	367
R^2	0.625	0.749	0.362	0.323
FE	Region	Region	—	Region

Heteroskedasticity-robust standard-errors in parentheses; col. (4) clustered by district
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Each column reports a separate regression of an intermediate-period outcome on Local Officials (%), the share of Ukrainian officials in 1897. Cols. (1)–(3) are at the historical-district level; col. (4) is at the Soviet raion level (367 raions matched to 84 imperial districts via ≥50% area overlap). Col. (1): land Gini coefficient in 1905. Col. (2): households employing six or more domestic servants per 1,000 population in 1897. Col. (3): share of urban population in 1926 (region FE omitted because $N = 81$ is too small). Col. (4): crude death rate per 1,000 population in 1929 (Harvard Ukrainian Research Institute historical GIS). Controls: standard geographic and historical set as in the notes to Table 2.

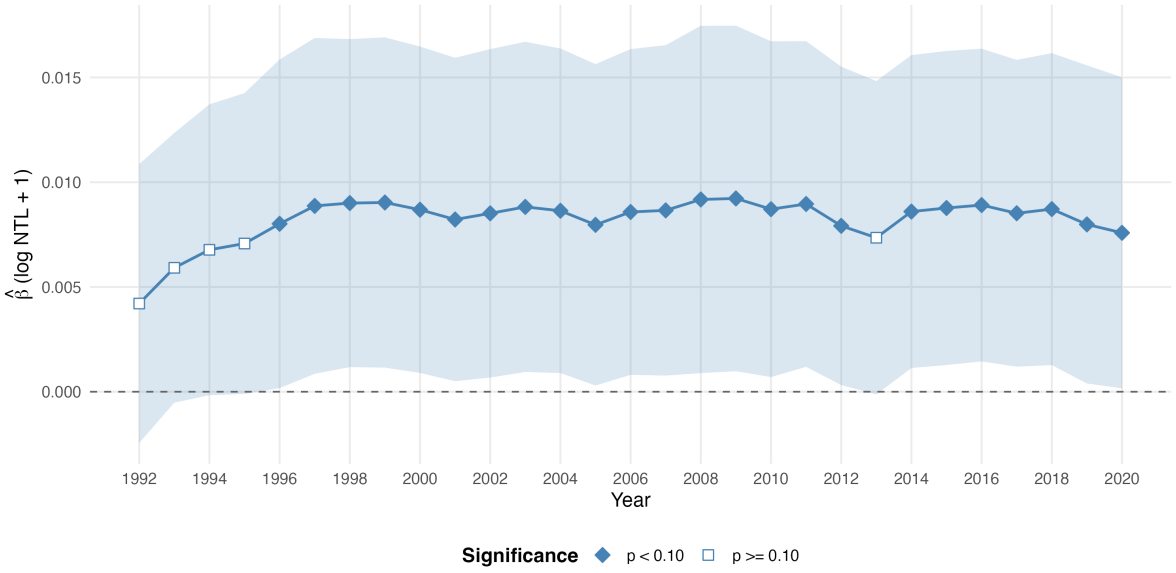

Figure C.1: Effect of historical self-governance on nighttime light intensity by year, 1992–2020

Notes: Each point represents the OLS coefficient on the share of local officials (%) from a separate regression of $\log(\text{NTL}+1)$ on the governance measure with geographic, soil, and historical controls plus region fixed effects ($N = 93$ districts, HC1 SEs). Diamonds indicate $p < 0.10$; squares indicate $p \geq 0.10$. Shaded area shows the 90% confidence interval. The LRCC-DVNL 2021 raster exhibits a distributional discontinuity inconsistent with the preceding series and is excluded.

Table C.2: Self-Governance, Collectivization, and Holodomor Mortality (Raion Level)

	Collectiviz.	Famine losses per 1000 (1933–34)			Procurement	Fulfillment
	rate (%)				plan 1932–33	rate (%)
	(1)	(2)	(3)	(4)	(5)	(6)
Share local officials (%)	0.2300**	0.8400	1.292*	0.5736	-1.353	-0.0320
	(0.1091)	(0.7527)	(0.7591)	(0.7857)	(0.9782)	(0.1200)
Rural pop. density (log)	-5.418	60.94***		67.22***		
	(3.403)	(15.71)		(14.88)		
Ukrainian pop. 1897 (%)	-0.0967	-1.157	-1.233	-1.045	2.007	0.0586
	(0.1328)	(0.9480)	(0.9706)	(0.9474)	(1.836)	(0.1656)
Ukrainian pop. 1926 (%)	0.1169*	0.6503**	0.4589	0.5149*	0.2410	-0.0185
	(0.0663)	(0.2850)	(0.2890)	(0.2678)	(1.103)	(0.0821)
Collectivization (%)			0.9978**	1.159**		
			(0.4576)	(0.4407)		
Observations	367	367	367	367	364	364
R^2	0.641	0.408	0.399	0.426	0.671	0.414
FE	Region	Region	Region	Region	Region	Region
District controls	✓	✓	✓	✓	✓	✓

Clustered (district) standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: OLS estimates at the raion level (367 Soviet raions matched to 84 imperial districts; each raion assigned to the district covering $\geq 50\%$ of its area). Column (1): percentage of households collectivized by late 1932. Columns (2)–(4): total famine losses per 1,000 population in 1933–34. Column (5): planned grain procurement in tons per 1,000 rural population, 1932–33. Column (6): grain procurement plan fulfillment rate (%), 1932–33. District-level controls include log population, geographic variables (river density, forest cover, slope, coastal distance, latitude, longitude, distance to Kyiv, chornozem share), and historical controls (Ukrainian population shares from 1897 and 1926, urbanization, merchants per 1,000, literacy, industrialization). Source: Harvard Ukrainian Research Institute (HURI) historical GIS.

Notes

1. I group the provinces into three macro-regions: Left Bank, Right Bank, and South, adapting traditional historiographic divisions. Sloboda Ukraine (Kharkiv province) is treated as part of the Left Bank: it was a frontier region settled by Cossacks who received limited self-governance privileges from Moscow in exchange for border defense. Volhynia belongs to the Right Bank. Western Ukraine, outside the Russian Empire, is excluded. Districts whose populations were predominantly non-Ukrainian (in present-day Russia, Belarus, and Moldova) are also excluded; an exception is the Putyvl district of Kursk province, which is assigned to the Left Bank because it lies almost entirely within modern Ukraine.

2. Chornozem (чорнозем) is the highly fertile black-earth soil that dominates central and southern Ukraine.

3. The sample is built from Ukraine's 1,469 hromadas: 199 communities in the western oblasts (Lviv, Ivano-Frankivsk, Zakarpattia) and 77 in Chernivtsi and Ternopil on the former Austro-Hungarian side are excluded for lack of Russian Empire district overlap; 31 in the occupied parts of Donbas are excluded for lack of fiscal data; 28 that do not map reliably to an 1897 district (17 with overlap below 50%, 11 with none), one covering only 3.2% of its matched district (Slovianoserbsk), and one with erroneous fiscal data (Pidhaitsi) are also dropped. This yields 1,132 communities.

4. Dropping log(distance to Kyiv) from the control vector, which is moderately correlated with the instrument ($r = 0.54$), raises the first-stage F from 11.8 to 19.5 and the IV point estimate from 0.017 to 0.025. The paper retains log(distance to Kyiv) as a control because

proximity to Kyiv plausibly affects contemporary local tax revenue through non-governance channels; the specification reported in the main table is therefore the more conservative choice.

5. Kursk is represented in the sample by a single district (Putyvl, in modern Sumy oblast). Province fixed effects include it as the eleventh level. Leave-one-out estimates cover only the ten provinces with multiple districts, since dropping Kursk would remove one district rather than a province-wide block.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, *117*(4), 1231–1294. https://doi.org/10.1162/003355302320935025

Acemoglu, D., & Robinson, J. A. (2012). *Why nations fail: The origins of power, prosperity and poverty*. Profile Books.

Becker, S. O., Boeckh, K., Hainz, C., & Woessmann, L. (2016). The Empire Is Dead, Long Live the Empire! Long-Run Persistence of Trust and Corruption in the Bureaucracy. *The Economic Journal*, *126*(590), 40–74.

Becker, S. O., & Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History*. *Quarterly Journal of Economics*, *124*(2), 531–596. https://doi.org/10.1162/qjec.2009.124.2.531

Bruhn, M., & Gallego, F. A. (2012). Good, Bad, and Ugly Colonial Activities: Do They Matter for Economic Development? *The Review of Economics and Statistics*, *94*(2), 433–461.

Buggle, J. C., & Nafziger, S. (2021). The Slow Road from Serfdom: Labor Coercion and Long-Run Development in the Former Russian Empire. *The Review of Economics and Statistics*, *103*(1), 1–17. https://doi.org/10.1162/rest_a_00857

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, *90*(3), 414–427. https://doi.org/10.1162/rest.90.3.414

Chaudhary, L., & Iyer, L. (2025). The Importance of Being Local? Administrative Decentralisation and Human Development [Advance access, article ueaf100]. *The Economic Journal*. https://doi.org/10.1093/ej/ueaf100

Cinelli, C., & Hazlett, C. (2020). Making Sense of Sensitivity: Extending Omitted Variable Bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(1), 39–67. https://doi.org/10.1111/rssb.12348

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*(1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0

Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly Exogenous. *Review of Economics and Statistics*, *94*(1), 260–272. https://doi.org/10.1162/REST_a_00139

Dell, M. (2010). The Persistent Effects of Peru's Mining *Mita*. *Econometrica*, *78*(6), 1863–1903. https://doi.org/10.3982/ECTA8121

Dell, M., Lane, N., & Querubin, P. (2018). The Historical State, Local Collective Action, and Economic Development in Vietnam. *Econometrica*, *86*(6), 2083–2121. https://doi.org/10.3982/ECTA15122

Giuliano, P., & Nunn, N. (2017, July). *Understanding Cultural Persistence and Change* (tech. rep. No. w23617). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w23617

Grosfeld, I., & Zhuravskaya, E. (2023). Persistent effects of empires: Evidence from the partitions of Poland. https://doi.org/halshs-00795231f

Harus, A., & Nivyevskyi, O. (2020). In Unity There Is Strength: The Effect of the Decentralization Reform on Local Budgets in Ukraine. *VoxUkraine*.

Hatsko, V., Brik, T., Tytiuk, S., Koval, A., Piddubnyi, I., Petrynka, K., Savisko, M., & Darkovich, A. (2025, April). KSE Local Data Hub. https://doi.org/10.5281/ZENODO.15267573

Iyer, L. (2010). Direct versus Indirect Colonial Rule in India: Long-Term Consequences. *The Review of Economics and Statistics*, *92*(4), 693–713. https://doi.org/10.1162/REST_a_00023

Kibita, N. (2013). Controlling Resources in the Ukrainian SSR during the Sovnarkhoz Reform (1957–1965): Testing the Idea of Decentralisation. *Europe-Asia Studies*, *65*(6), 1171–1191. https://doi.org/10.1080/09668136.2013.806136

Markevich, A., Naumenko, N., & Qian, N. (2025). The Causes of Ukrainian Famine Mortality, 1932–33. *Review of Economic Studies*, *92*(5), 3276–3305. https://doi.org/10.1093/restud/rdae091

Markevich, A., & Zhuravskaya, E. (2018). The Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire. *American Economic Review*, *108*(4-5), 1074–1117. https://doi.org/10.1257/aer.20160144

Michalopoulos, S., & Papaioannou, E. (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, *81*(1), 113–152. https://doi.org/10.3982/ECTA9613

Peisakhin, L. (2013). Long-Run Persistence of Political Attitudes and Behavior: A Focus on Mechanisms.

Pop-Eleches, G. (2015). Pre-Communist and Communist Developmental Legacies. *East European Politics and Societies: and Cultures*, *29*(2), 391–408. https://doi.org/10.1177/0888325414555761

Romanova, V., & Umland, A. (2019). Ukrainian Local Governance Prior to Euromaidan: The Pre-History of Ukraine's Decentralization Reform. *Kennan Cable*, *44*.

Sokoloff, K. L., & Engerman, S. L. (2000). Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives*, *14*(3), 217–232. https://doi.org/10.1257/jep.14.3.217

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, *65*(3), 557–586. https://doi.org/10.2307/2171753

Tang, H., Zhong, Y., & Xia, H. (2025, April). Global nighttime light dataset from 1992 to 2022 with focus on low-light areas. https://doi.org/10.7910/DVN/15IKI5

Trojnitskij, N. (1903—1905). *Pervaya Vseobshchaya Perepis' Naseleniya Rossijskoj Imperii 1897 Goda. (Gubernskie Itogi) [Первая всеобщая перепись населения Российской Империи 1897 Года (Губернские Итоги)]* (T. 1–89). Statotdel TsSK press.

U-LEAD. (2025, September). Byudzhety Terytorialnykh Hromad Ukrayiny [Detsentralizatsiya v Ukrayini Бюджети Територіальних Громад України].

WorldPop & Bondarenko, M. (2020). Individual Countries 1km UN Adjusted Population Density (2000-2020). https://doi.org/10.5258/SOTON/WP00675

Zhao, Q. C. (2023, June). Rethinking "Distance From": Lessons from Wittenberg and Mainz.

Online Appendix to "Local Officials in Imperial Administration and Long-Run Development in Ukraine"

Marian Melnyk

June 2, 2026

Online Appendix A documents the historical background, the 2014 decentralization reform, contemporary data and sample-matching procedure, and definitions of the outcome and fiscal-autonomy measures. Online Appendix B extends the body's nighttime-lights persistence finding to the ATC level and to a Crimea-inclusive district sample, and reports two additional mechanism tests on Soviet-era physical capital and late-imperial human capital. Online Appendix C reports Conley spatial standard errors and a wild cluster bootstrap, plausibly exogenous IV bounds, sensitivity to unobservable confounding, and a capital-city versus rural decomposition of the treatment.

Online Appendix A. Methodological Detail

Historical Background

In 1648, Bohdan Khmelnytskyi led a Cossack uprising against the Polish-Lithuanian Commonwealth that, at its peak, covered large parts of present-day central, northern, eastern, and southwestern Ukraine on both banks of the Dnipro. The revolt established the Cossack Hetmanate, a semi-autonomous polity with its own military and administrative system. However, it faced constant threats from Polish counteroffensives and limited external recognition. Seeking security, Khmelnytskyi sought military protection from Muscovy and the Crimean Khanate. The Pereyaslav Agreement of 1654 aligned the Hetmanate with Muscovy in exchange for military protection, while preserving broad autonomy, including the right to elect its leaders. Over time, however, this relationship gradually shifted the balance of power toward Muscovy, eroding the Hetmanate's autonomy and deepening imperial control (Hajda et al., 2025).

Continued conflict between Muscovy and the Commonwealth culminated in the Truce of Andrusovo (1667), which divided Ukraine along the Dnipro River, returning the Right Bank to Polish control and interrupting the development of autonomous institutions there. The Hetmanate retained diminished autonomy under Muscovy, but successive centralizing reforms gradually curtailed its independence. Rebellions periodically arose to restore wider autonomy, the most prominent being Hetman Ivan Mazepa's alliance with Swedish King Charles XII during the Great Northern War. After Mazepa's defeat at Poltava in 1709, Russian forces destroyed his capital, Baturyn, as punishment for his defection. Baturyn's historical role as the Hetmanate's administrative center motivates the instrumental variable strategy described in Section 2.

In 1721, the Tsardom of Muscovy was formally proclaimed the Russian Empire. The Hetmanate continued to exist in diminished form for another half-century within this new imperial framework, but its administrative independence was steadily eroded. Under Kyrylo

Rozumovskyi, the last hetman, Baturyn was briefly revived as the administrative capital but never regained its former political or economic significance. Finally, in 1764, Empress Catherine II abolished the office of hetman and replaced the Hetmanate's structures with a new imperial administration, formally ending Cossack autonomy. A year later, a similar process unfolded in Sloboda Ukraine, where Catherine revoked local privileges and integrated the region under direct imperial governance.

Further imperial expansion came in 1793 with the Second Partition of Poland, which brought Right-bank Ukraine under Russian rule. In contrast to the Left Bank, these territories lacked sustained autonomous institutions. Governance traditions that existed during the Hetmanate era were interrupted by the 1667 partition and subsequent Polish and then Russian direct rule. By the late eighteenth century, the empire had also consolidated control over the South: Crimea was annexed in 1783, territories between the Southern Buh and Dnister Rivers in 1792, and Bessarabia (the southern part of which lies in modern Odesa oblast) in 1812 following Russia's victory over the Ottoman Empire.

Historical Data Sources

The study focuses on districts, historically known as *uezd*, a level-3 subnational administrative unit in the Russian Empire (also known as *povit*). One hundred and two povits fall within modern Ukraine's borders; the main estimation sample is the 93 with modern ATC coverage (excluding seven Crimean povits, Slovianoserbsk, and Taganrog).

New historical datasets were collected from the censuses and historical statistical yearbooks from the late nineteenth century and merged with contemporary data. The share of local officials in district administration, police, and courts is a continuous measure of historical governance experience in 1897. However, the main source (Trojnitskij, 1903—1905) lacks information for Hlukhiv, the capital town of Hlukhiv district, because several pages are missing from the scanned volumes. To fill the gap, I impute the count of Ukrainian officials in Hlukhiv town with the average across other district-capital cities in Chernihiv province.

Of the fourteen other district capitals, I exclude four (Mhlyn, Novozybkiv, Starodub, and Surazh) whose cities had Ukrainian population shares below 1% and which lie beyond modern Ukraine's borders, and I exclude the provincial capital Chernihiv due to its disproportionate economic and political weight. The averaging pool is therefore the nine remaining district capitals (Nizhyn, Konotop, Borzna, Krolevets, Sosnytsia, Novgorod-Seversky, Oster, Kozelets, and Horodnia), whose districts each had an aggregate Ukrainian population share above 86 percent, comparable to Hlukhiv district (91.6%). The source data records officials separately for each town status (rural vs. capital city); because only the Hlukhiv-city pages are missing, the imputation augments the district numerator (the count of Ukrainian officials) without altering the district denominator. Seven modern communities map to Hlukhiv district and are therefore affected by this imputation.

Yasnopolskij (1913) includes information on the number of factories, factory workers, production, and population for 1907-1909 at the district-level. I use it to define industrialization as the number of factories per 10,000 people. Southern provinces had considerably more factories per 10,000 people than other regions, with the Right-bank provinces being the least industrialized. However, the book lacks district-level data for Bessarabia province. Unlike other provinces, where data are available by district, the book reports Bessarabia only at the province level; the author notes that manufacturing in Bessarabia was underdeveloped. Additionally, the administrative complexities in Bessarabia, particularly in the Izmail and Akkerman districts, which were transferred between the Moldavian Principality and the Russian Empire under the Treaty of Paris (1856) and the Berlin Treaty (1878), further complicated data collection. The distinct internal structure of the Izmail district, which remained in place until it was fully integrated into the Russian Empire in 1904, might have made consistent reporting across districts challenging even after this period.

To address this, I redistributed the 227 factories across the eight districts using a weighted distribution method, first calculating the total population of each district and then assigning factories based on their share of the province's population. The production data was

distributed similarly, though data on the number of workers is missing even at the province level. This adjustment affects only three districts, as only parts of Bessarabia fall within Ukraine's modern borders. In a subsample analysis discussed later, data from this province are entirely excluded, providing an additional robustness check.

The 2014 Decentralization Reform

The reform that created the sample hromadas (the Ukrainian-language term for ATCs, used interchangeably throughout) ran from 2014 to 2020 and transferred substantial fiscal and administrative powers to newly amalgamated communities. A concept paper was published on 1 April 2014; on 28 December 2014, the Parliament amended the Budget Code to entitle voluntarily amalgamated villages and cities to direct transfers from the Ministry of Finance, and a Law on voluntary amalgamation followed on 5 February 2015 (Arends et al., 2023). Over six years, roughly 10,000 local councils merged into 1,469 new hromadas; the process became routine after 2017 and closed in October 2020, with communities that had not voluntarily amalgamated consolidated by central directive.

Communities received authority over communal property, housing and utilities, local transport, primary and secondary education, primary and (where available) secondary health care, libraries and childcare, fire and civil protection, roads of local significance, and land-use planning; from 2018 they also received ownership of formerly state-owned land within their jurisdictions (Arends et al., 2023; Rabinovych et al., 2024). Salaries for schoolteachers and medical personnel remained a central-government responsibility, but hromadas gained the right to select school directors and reorganize the local school network (Arends et al., 2023).

The fiscal core of the reform was a jump in the personal income tax (PIT) retained by local budgets from 25 to 60 percent in the new hromadas (64 percent from 2022), combined with the locally set taxes and fees described in the Local Tax Revenues subsection below (Rabinovych et al., 2024). One institutional feature matters for the empirical measures used in the paper: PIT is allocated mechanically to the hromada where an employer is

registered, not where the employee lives or the enterprise physically operates (Rabinovych et al., 2024). This wedge between governance effort and PIT revenue motivates the "local tax share excluding transfers" measure introduced later, which nets the mechanically-allocated PIT share out of the denominator.

Contemporary Data

Contemporary data come from Ukraine's 1,469 Amalgamated Territorial Communities (ATCs), established under the decentralization reform initiated in 2014. The reform transferred fiscal and administrative powers from higher authorities to local communities, boosting tax revenues, spending efficiency, and satisfaction with public services (Harus & Nivyevskyi, 2020; Reznik et al., 2020) and strengthening local trust (Arends et al., 2023). This created a setting in which local governance quality and administrative competence can shape economic outcomes.

Complete data are available only from 2021 onward, as community formation was finalized in 2020; communities that had not amalgamated voluntarily by that date were consolidated through top-down directives. Post-2021 data reflect the disruptions of the ongoing war (Rabinovych et al., 2024).

The full sample excludes the territories that were not part of the Russian Empire. Additionally, community data are not available for the annexed Crimean peninsula. There are also 31 communities in the Donbas region (Donetsk and Luhansk oblasts) that are temporarily occupied, and so community data do not exist for them either.

Only one community has errors in the reported data, and so I dropped it from the sample (Pidhaitsi community in Volyn oblast). Table OA.A.3 shows the distribution of the 1,132 sample communities by size category.

Matching Modern and Historical Boundaries

Ukraine comprises 1,469 territorial communities (*hromady*) established by the 2020 administrative reform; Crimea was not divided into hromady due to Russian occupation. The estimation sample is restricted to communities whose population-weighted territory falls predominantly within Russian Empire districts. Communities in three western oblasts that were historically part of the Austro-Hungarian Empire (Lviv, Ivano-Frankivsk, and Zakarpattia, with 199 communities) are excluded because they have no overlap with the Russian Empire district boundaries. Chernivtsi and Ternopil oblasts straddle the historical border between the Russian and Austro-Hungarian empires: 77 communities whose territory primarily belonged to the Austro-Hungarian side are excluded for the same reason, while the 30 communities on the Russian Empire side are retained at this step. A further 31 communities in the temporarily occupied parts of Donetsk and Luhansk oblasts are excluded because contemporary fiscal data are not available; the 67 communities in government-controlled Donbas remain in the sample.

The estimation sample further excludes 28 hromadas that do not map reliably to Russian Empire districts: 17 whose best-matching district overlap falls below 50% (Table OA.A.1) and 11 whose territory has no overlap with any 1897 district boundary. One additional community (Slovianoserbsk, Luhansk oblast) is excluded because its single ATC covers only 3.2% of the matched historical district. The resulting sample is 1,132 communities.

Table OA.A.1: Hromadas excluded from the estimation sample (largest 1897-district overlap $< 50\%$)

Oblast	Raion	Community	Assigned District	Province	Ratio (%)
Ternopil	Chortkiv	Skala-Podilska	Kamianets-Podilskyi	Podilsk	0.03
Ternopil	Chortkiv	Husiatyn	Kamianets-Podilskyi	Podilsk	0.36
Ternopil	Ternopil	Pidvolochysk	Starokostiantyniv	Volhynia	0.50
Chernivtsi	Chernivtsi	Vikno	Khotyn	Bessarabia	0.61
Ternopil	Ternopil	Skorykiv	Kremenets	Volhynia	1.85
Ternopil	Chortkiv	Hrymailiv	Proskuriv	Podilsk	2.78
Ternopil	Ternopil	Zbarazh	Kremenets	Volhynia	19.47
Chernivtsi	Chernivtsi	Yurkivtsi	Khotyn	Bessarabia	21.59
Ternopil	Chortkiv	Melnytsia-Podilska	Khotyn	Bessarabia	29.01
Mykolaiv	Mykolayiv	Ochakiv	Odesa	Kherson	29.04
Chernivtsi	Chernivtsi	Toporivtsi	Khotyn	Bessarabia	35.51
Chernivtsi	Chernivtsi	Novoselytsia	Khotyn	Bessarabia	36.69
Poltava	Poltava	Nekhvoroscha	Kostiantynohrad	Poltava	37.16
Khmelnytsk	Khmelnytsk	Staryi Ostropil	Novohrad-Volynskyi	Volhynia	40.19
Kyiv	Brovary	Zhurivka	Pryluky	Poltava	42.06
Odesa	Podilsk	Okny	Ananiv	Kherson	42.11
Odesa	Berezivka	Berezivka	Tyraspol	Kherson	43.17

Notes: Hromadas where the population-weighted share of the 2020 territory falling into the single best-matching 1897 Russian Empire district is below 50%. These 17 communities are excluded from the main N=1,132 estimation sample. Ratio is the population share (%) overlapping the named district.

Summary Statistics

Table OA.A.2: Summary Statistics

						Mean by Local Officials	
Variable	*N*	Mean	SD	Min	Max	Below median	Above median
Panel A: Outcome and key explanatory variable							
Local tax revenue p.c. (UAH)	1,132	2,127	1,572	349	20,733	2,059	2,192
Share of local officials (1897, %)	1,132	44.54	20.01	11.56	95.68	27.85	60.82
Panel B: Geographic controls							
Population density	1,132	0.14	0.57	0.01	10.44	0.17	0.11
Population	1,132	25,678	78,583	1,814	1,511,554	29,227	22,217
Elevation (m)	1,132	150	70.05	0.01	346	142	158
Terrain ruggedness	1,132	36.95	20.32	7.13	153	36.76	37.14
Slope (mean)	1,132	0.62	0.37	0.05	2.94	0.59	0.65
River density	1,132	0.23	0.06	0.00	0.86	0.23	0.23
Forest cover	1,132	0.21	0.17	0.00	0.85	0.19	0.22
Chernozem share	1,132	0.66	0.42	0.00	1.00	0.63	0.69
Log dist. to coast	1,132	5.27	1.16	-0.50	6.61	4.94	5.58
Log dist. to Baturyn	1,132	12.72	0.57	8.16	13.51	13.01	12.44
Longitude	1,132	31.56	3.63	23.80	39.97	30.61	32.50
Latitude	1,132	49.11	1.51	45.36	52.24	48.74	49.47
Panel C: Historical and demographic controls							
Ukrainian pop. share (%)	1,132	72.08	18.73	19.60	98.10	58.19	85.63
Urbanization rate (1897, %)	1,132	13.86	7.23	5.98	28.86	15.92	11.86
Merchants per 1,000 (1897)	1,132	2.82	2.29	0.11	10.55	3.49	2.17
Literacy rate (1897, %)	1,132	19.70	7.05	7.10	43.94	23.32	16.17
Factories per 10,000 (1908)	1,132	1.86	1.98	0.02	7.92	1.91	1.81
Panel D: Fiscal capacity outcomes (2021)							
Subsidy dependence rate	1,132	0.06	0.06	0.00	0.30	0.08	0.05
Transfer share	1,132	0.42	0.13	0.08	0.83	0.47	0.39
Local tax share (excl. transfers)	1,132	0.38	0.11	0.04	0.86	0.39	0.37
Own-source revenue p.c. (UAH)	1,132	6,606	4,206	1,488	58,866	6,076	7,123
Culture & sport spending p.c. (UAH)	1,132	486	260	26.59	3,694	453	518

Notes: The sample consists of 1,132 communities (ATCs) matched to 93 historical districts of the Russian Empire. The last two columns report means for communities in districts with below- vs. above-median shares of local officials in 1897 (median = 40.6%). Below median: $N = 559$; above median: $N = 573$. Variables entering regressions in logs are shown in levels.

Table OA.A.3: Distribution of ATCs by size type

	Size type	Count
1	oblast center	13
2	more than 30,000	166
3	15,000 to 30,000	245
4	10,000 to 15,000	188
5	5,000 to 10,000	352
6	less than 5,000	168

Notes: Distribution of the 1,132 ATCs (2020 hromada boundaries, main estimation sample) across six size categories.

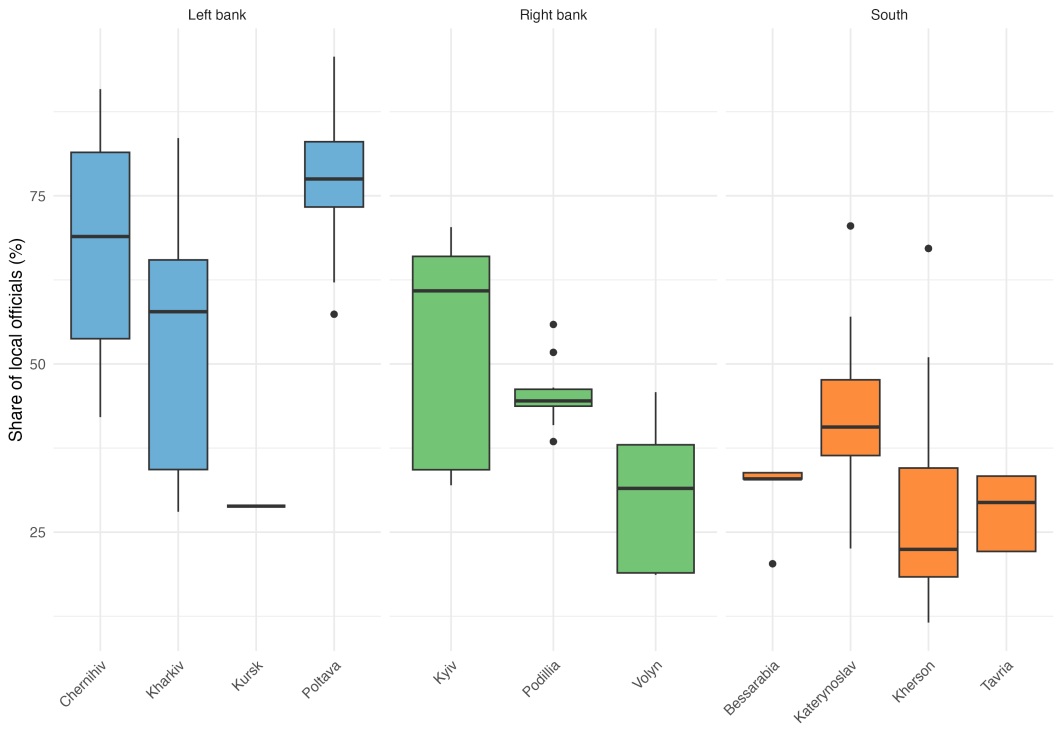

Figure OA.A.1: Distribution of local officials' share by province and macro-region

Local Tax Revenues as the Main Outcome Variable

The main outcome variable is per-capita local tax revenue in 2021, expressed in nominal Ukrainian hryvnia (UAH) and drawn from the U-LEAD Budgets of Territorial Communities (U-LEAD, 2025), which compiles hromada-level execution reports submitted by local treasuries to Ukraine's Ministry of Finance. Revenues are divided by the resident population used by the source to compute per-capita values; no deflator is applied because the analysis is cross-sectional in a single fiscal year. Across the 1,132 sample communities, local tax revenue averages 2,127 UAH per capita (SD 1,572; range 349–20,733), equivalent to roughly 78 USD at the 2021 average exchange rate of 27.3 UAH per US dollar. Local taxes and fees form the foundation of communities' fiscal autonomy under the post-2014 framework; across sample hromadas, they average 38 percent of own-source revenue (Table OA.A.2, Panel D).

According to Article 10 of the Tax Code of Ukraine, local taxes and fees are established directly by village, settlement, or city councils within the limits defined by the Code. These include two broad categories:

1. Local taxes: (i) the property tax, which itself comprises the land tax, real estate tax (on property other than land), and transport tax; and (ii) the single tax applied to small businesses under a simplified taxation regime.

2. Local fees: the parking fee and the tourist fee.

The establishment of any other taxes or levies not specified in the Tax Code is prohibited, ensuring a uniform nationwide framework while leaving meaningful fiscal discretion at the local level through rate-setting, exemptions, and administration.

Collection and administration are divided: the State Tax Service of Ukraine, operating through its oblast and raion offices, registers taxpayers and enforces payment, while hromada councils set rates, grant exemptions, and bear the fiscal consequences of under- or over-collection. Rate-setting is meaningful in practice (Arends et al., 2023). Communities exercise discretion within the Tax Code bands, particularly on real estate, land, and single-tax Group

I and II rates, and the resulting variation in revenue per capita across hromadas reflects both the local economic base and the active fiscal choices of the council.

The *single tax* is an alternative to the general taxation system designed to reduce administrative burden and incentivize small-scale entrepreneurship. It applies to four groups of taxpayers, classified by activity type, income limits, and employment scale:

- Group I – self-employed individuals engaged in minor retail or service activities without hired workers with annual revenue $\leq$ 300,000 UAH. Local councils set the rate at up to 10% of the subsistence minimum for able-bodied persons in force on January 1 of the tax year (2,270 UAH per month as of January 1, 2021), implying a maximum monthly single-tax payment of 227 UAH.

- Group II – small entrepreneurs with a few employees ($\leq$ 10) with annual revenue $\leq$ 1.5 M UAH. Local councils set the rate at up to 20% of the minimum wage in force on January 1 of the tax year (6,000 UAH per month as of January 1, 2021), implying a maximum monthly payment of 1,200 UAH.

- Group III – both individuals and legal entities with annual revenue $\leq$ 5 M UAH; they pay a share of quarterly revenue (3% if registered for VAT, 5% if not).

- Group IV – agricultural producers taxed based on the area and category of agricultural land used.

Local governments have the authority to set specific rates (within national limits) and thus directly influence the level of revenue from this tax. Importantly, payers of the single tax are exempt from several national taxes, including the corporate profit tax, personal income tax (on business income), and the land tax (for land used in business). This makes the single tax an especially critical component of local budgets, as it concentrates a large share of entrepreneurial activity within local authorities' jurisdictions.

In this study, local taxes are measured as per-capita revenues from own sources (aggregated across all local taxes and fees). Because revenue levels reflect both governance

effort and the local economic base, the next subsection introduces ratio-based measures that partially net out the base.

The outcome year 2021 is the earliest year with uniform hromada-level coverage following the 2020 consolidation and the last year fully predating the February 2022 invasion. It is also the second year of the COVID-19 pandemic, which depressed parking and tourist fees and, more broadly, the activity of small entrepreneurs on the simplified regime; these effects fall on levels and should not bias cross-sectional comparisons across hromadas within the same fiscal year.

Fiscal Autonomy Measures

To examine whether the main result reflects broader fiscal autonomy rather than a single outcome, I draw on the KSE Local Data Hub (Hatsko et al., 2025), a repository of hromada-level indicators compiled from official budget reporting by researchers at the Kyiv School of Economics (see also Rabinovych et al., 2024) that covers 1,438 communities, including the full 1,132-community estimation sample. Under the post-2014 fiscal framework, communities' operating budgets (the "general fund") are financed from four sources: locally set taxes and fees, a retained share of the personal income tax (PIT, increased from 25% to 60% of PIT collected within the community's boundaries (Arends et al., 2023)), a retained share of excise tax on retail sales of fuel, alcohol, and tobacco, and intergovernmental transfers, which include both equalization grants and earmarked subventions (targeted transfers for specific purposes such as education or healthcare). Under the equalization system, the central government provides grants to communities whose per capita tax capacity falls below a norm and collects reverse transfers from those above it.

I use four measures that capture distinct dimensions of fiscal autonomy: transfer dependence (measures 1–2), the composition of own-source revenue (measure 3), and discretionary spending power (measure 4). Focusing on composition rather than levels follows recent evidence that own-revenue share predicts hromada institutional performance (preparedness to

wartime shocks) more strongly than total revenue per capita (Rabinovych et al., 2024). The measures are defined as follows:

1. The *subsidy dependence rate*: equalization grants as a share of general fund revenue. Negative values indicate net contributors to the equalization system, i.e., communities whose tax base exceeds the norm.

2. The *transfer share*: all intergovernmental transfers as a proportion of total budget revenue.

3. The *local tax share excluding transfers*: local taxes and fees divided by own-source revenue (general fund revenue minus all transfers). The denominator thus comprises locally set taxes, the retained 60% PIT share, and excise taxes. A higher ratio indicates greater reliance on taxes that the community actively administers (such as property and business taxes), relative to the PIT share, which is allocated mechanically based on where employers are registered rather than on community governance effort.

4. *Per capita spending on culture, physical education, and sport*: unlike education and health expenditures, which are subject to central norms and partially funded by earmarked transfers, spending on culture and sport is financed entirely from communities' own revenues and reflects local prioritization.

Data Sources

Tables OA.A.4–OA.A.6 catalog every variable used in the analysis, grouped by outcomes, treatment, and instrument (Table OA.A.4); historical, geographic, and soil controls (Table OA.A.5); and boundaries, population weights, and mechanism datasets (Table OA.A.6). Each entry lists the variable, its level and time period, and the source.

Table OA.A.4: Data sources — outcomes, treatment, and instrument

Variable	Level	Year(s)	Source
Panel A. Outcome variables			
Local tax per capita	Hromada	2021	U-LEAD (2025)
Subsidy dependence	Hromada	2021	KSE Local Data Hub (Hatsko et al., 2025)
Transfer share	Hromada	2021	KSE Local Data Hub (Hatsko et al., 2025)
Local tax share excl. transfers	Hromada	2021	KSE Local Data Hub (Hatsko et al., 2025)
Culture/sport spending pc	Hromada	2021	KSE Local Data Hub (Hatsko et al., 2025)
Nighttime light intensity	District	1992–2020	LRCC-DVNL (Tang et al., 2025)
Panel B. Treatment and instrument			
Share of local officials (%)	District	1897	1897 Imperial Census (Trojnitskij, 1903—1905)
Log distance to Baturyn	Hromada	time-invariant	Author's calculation; centroid distance to former Hetmanate capital

Notes: Outcome and treatment variables enter the main estimating sample of 1,132 hromadas across 93 historical districts. Per-capita variables use the population reported by each source. The instrument is the great-circle distance from each hromada centroid to Baturyn, the former Hetmanate capital.

Table OA.A.5: Data sources — historical, geographic, and soil controls

Variable	Level	Year(s)	Source
Panel C. Historical controls			
Ukrainian population share	District	1897	1897 Imperial Census (Trojnitskij, 1903—1905)
Urbanization rate	Gubernia	1897	Markevich et al. (2025) replication data
Literacy	District	1897	1897 Imperial Census (Trojnitskij, 1903—1905)
Merchants per 1,000 pop.	District	1897	1897 Imperial Census (Trojnitskij, 1903—1905)
Elite servants per 1,000 pop.	District	1897	1897 Imperial Census (Trojnitskij, 1903—1905)
Industrialization (factories/10k)	District	1907–1909	Yasnopolskij (1913)
Panel D. Geographic and soil controls			
Mean slope (degrees)	Hromada / District	time-invariant	SRTM 30 arc-second elevation (Jarvis et al., 2008)
River density (km/km^2)	Hromada / District	time-invariant	HydroRIVERS v1.0 (Lehner & Grill, 2013)
Forest cover (%)	Hromada / District	2020	ESA WorldCover (Zanaga et al., 2021)
Chernozem soil share	Hromada / District	time-invariant	FAO Digital Soil Map of the World (Food and Agriculture Organization of the United Nations, 2003)
Log distance to Black Sea	Hromada / District	time-invariant	Natural Earth 10 m coastline (Kelso & Patterson, 2010)
Log distance to Kyiv	Hromada / District	time-invariant	Author's calculation from polygon centroids
Centroid longitude, latitude	Hromada / District	time-invariant	Polygon centroids of hromada/district geometries

Notes: District-level historical variables are assigned to hromadas via the population-weighted matching procedure described above. Distance to Kyiv is the great-circle distance from each hromada centroid to Kyiv (50.45°N, 30.52°E).

Table OA.A.6: Data sources — boundaries, population weights, and mechanism datasets

Variable	Level	Year(s)	Source
Panel E. Boundaries and population weights			
1897 district boundaries	District	1897	RISTAT Russian Empire Historical GIS (Kessler, 2025)
Hromada boundaries (1,469)	Hromada	2020	Ukrainian decentralization-reform administrative records
Population raster (1 km)	Grid	2020	WorldPop, used for population-weighted hromada→district matching (WorldPop & Bondarenko, 2020)
Panel F. Mechanism and robustness datasets			
Land Gini coefficient	District	1905	Buggle and Nafziger (2021) replication data
Schools per 1,000 pop.	District	1911	Buggle and Nafziger (2021) replication data
Soviet road density	District	early 1990s	Buggle and Nafziger (2021) replication data
Urbanization rate	District	1926	Markevich et al. (2025) replication data
Crude death rate per 1,000	Raion	1929	HURI MAPA Digital Atlas of Ukraine (Harvard Ukrainian Research Institute, 2024)
Installed electrical power	District (via oblast crosswalk)	1928–1969	Naumenko (2026) replication data
Collectivization rate (Oct–Dec)	Raion	1932	HURI MAPA (Harvard Ukrainian Research Institute, 2024)
Grain procurement / fulfillment	Raion	1930–1933	HURI MAPA (Harvard Ukrainian Research Institute, 2024)
Excess mortality / famine losses	Raion / District	1932–1934	HURI MAPA (Harvard Ukrainian Research Institute, 2024); Markevich et al. (2025)
Ukrainian population share	Raion	1926	HURI MAPA (Harvard Ukrainian Research Institute, 2024); areally interpolated to 1932 raion grid

Notes: Modern hromada population is taken from 2020 administrative records reported in the hromada boundary file. Mechanism and robustness datasets in Panel F enter only the corresponding appendix tables (electrification, schools, collectivization, famine mortality, persistence regressions).

Online Appendix B. Additional Persistence and Mechanism Tests

The headline district-level NTL persistence finding in body Appendix C, Figure C.1 is positive and marginally significant across nearly all years between 1992 and 2020.[1] This section extends that result in two directions. The first two figures report robustness checks on the NTL series: an ATC-level version of the regression for 2014–2020 (Figure OA.B.1) and a district-level version on an extended sample that adds Crimea and Slovianoserbsk (Figure OA.B.2). The remaining two tables test channels through which historical self-governance could have persisted: Soviet-era physical capital (Table OA.B.1) and late-imperial human capital (Table OA.B.2).

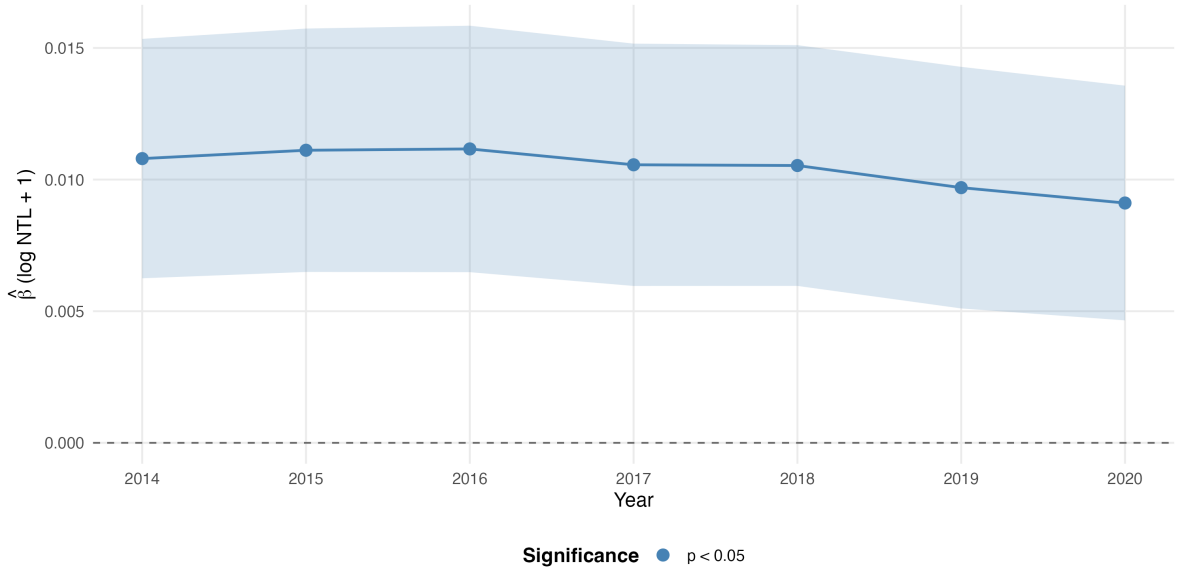

Figure OA.B.1: Effect of historical self-governance on nighttime light intensity by year, ATC level, 2014–2020

Notes: Each point represents the OLS coefficient on the share of local officials (%) from a separate regression of log(NTL + 1) on the governance measure with demographic, geographic, soil, and historical controls plus region fixed effects ($N = 1{,}132$ ATCs, SEs clustered by district). Shaded area shows the 90% confidence interval.

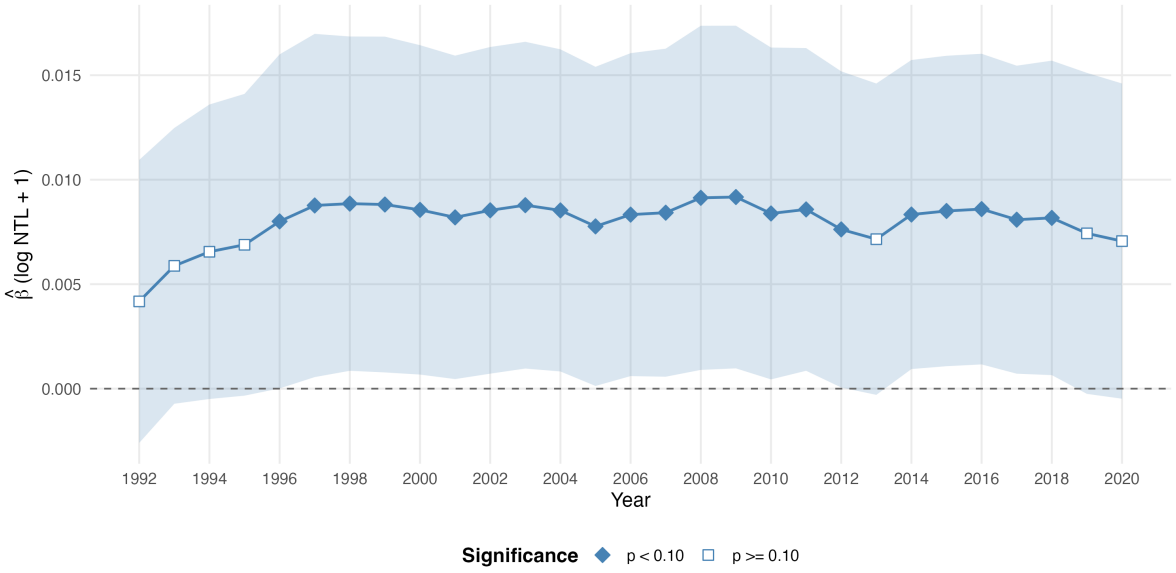

Figure OA.B.2: Effect of historical self-governance on nighttime light intensity by year, extended district sample, 1992–2020

Notes: Robustness to the main district-level specification in Figure C.1. The extended sample adds five Crimean povits (Perekop, Yevpatoria, Yalta, Simferopol, Feodosia) and Slovianoserbsk to the 93-district main sample ($N = 99$). Kerch-Yenikalsk and Sevastopol are excluded as city-only governorates (градоначальства); Taganrog is excluded because the district was administratively part of the Don Cossack Host rather than a Ukrainian province, and its capital city lies in modern Russia. Each point represents the OLS coefficient on the share of local officials (%) from a separate regression of $\log(\text{NTL} + 1)$ on the governance measure with geographic, soil, and historical controls plus region fixed effects (HC1 standard errors). Diamonds indicate $p < 0.10$; squares indicate $p \geq 0.10$. Shaded area shows the 90% confidence interval. Coefficients remain positive in every year and marginally significant ($p < 0.10$) in 22 of the 29 years, with magnitudes essentially identical to the main specification.

Table OA.B.1: Self-Governance and Soviet-Era Physical Capital

	log(Electricity capacity)					Road
	1928	1932	1946	1959	1969	density
	(1)	(2)	(3)	(4)	(5)	(6)
Local Officials (%)	-0.002	-0.007	0.002	-0.005	0.001	7.87×10^{-5}
	(0.013)	(0.014)	(0.006)	(0.004)	(0.006)	(7.12×10^{-5})
Observations	89	89	93	93	93	93
R^2	0.725	0.813	0.814	0.884	0.826	0.444
FE	Region	Region	Region	Region	Region	Region

Heteroskedasticity-robust standard-errors in parentheses

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Each column reports a separate district-level regression of a Soviet-era physical capital outcome on Local Officials (%), the share of Ukrainian officials in 1897. Cols. (1)–(5): log of installed electrical generating capacity in 1928, 1932, 1946, 1959, and 1969, drawn from Naumenko (2026). Sample size varies because some districts lack early-period electrification data. Col. (6): log road density in the early 1990s, drawn from Buggle and Nafziger (2021). Controls: standard geographic and historical set as in the notes to Table 2. Region fixed effects included throughout.

Table OA.B.2: Historical Self-Governance and Human Capital

Model:	Literacy 1897 (OLS) (1)	Literacy 1897 (IV) (2)	Schools 1911 (OLS) (3)	Schools 1911 (IV) (4)
Variables				
Local Officials	-0.0004	0.0036	0.0034	0.0034
	(0.0003)	(0.0027)	(0.0048)	(0.0285)
FE	Region	Region	Region	Region
Fit statistics				
Observations	93	93	93	93
R^2	0.86661	0.87093	0.69701	0.69513

Heteroskedasticity-robust standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$
Notes: District-level sample. Dependent variables: literacy rate in 1897 (cols 1–2) and schools per 1,000 in 1911 (cols 3–4). All specifications include geographic, soil, and historical controls with macro-region fixed effects; IV columns instrument Local Officials with log distance to Baturyn. HC1 robust standard errors.

Online Appendix C. Robustness and Sensitivity Analysis

Alternative Inference

Two concerns motivate alternative inference for the main estimates. First, observations within historical districts may be spatially correlated beyond what cluster-robust standard errors allow. Table OA.C.1 reports Conley (1999) spatial standard errors at distance cutoffs of 50, 100, and 200 km; OLS and IV point estimates retain statistical significance across all cutoffs and all fixed-effect specifications. Second, with 93 historical clusters, asymptotic cluster-robust inference may be unreliable. Table OA.C.2 reports a wild cluster bootstrap confidence interval for the preferred OLS specification and an Anderson–Rubin confidence interval for the IV; both leave the main conclusions unchanged.

Table OA.C.1: Robustness: Conley (1999) Spatial Standard Errors at Multiple Cutoffs

		Conley cutoff		
Spec.	FE	50 km	100 km	200 km
OLS	—	0.0077***	0.0077***	0.0077*
		(0.0019)	(0.0022)	(0.0040)
OLS	Region	0.0056***	0.0056**	0.0056**
		(0.0019)	(0.0023)	(0.0028)
OLS	Province	0.0071***	0.0071*	0.0071*
		(0.0026)	(0.0041)	(0.0041)
IV	—	0.0165***	0.0165***	0.0165***
		(0.0049)	(0.0044)	(0.0061)
	Wald F (1st stage)	11.57	9.00	8.39
IV	Region	0.0199*	0.0199***	0.0199***
		(0.0102)	(0.0044)	(0.0065)
	Wald F (1st stage)	4.63	4.06	3.80
IV	Province	0.0309**	0.0309**	0.0309**
		(0.0155)	(0.0137)	(0.0140)
	Wald F (1st stage)	4.55	4.59	6.43
Observations		1,132		
Controls		Yes (15 covariates)		

Conley (1999) spatial standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Dependent variable: log local tax revenue per capita (UAH, 2021). Sample: 1,132 ATCs. Specification matches Table 2 (full controls; OLS uses Region or Province FE as indicated; IV instruments Local Officials with log distance to Baturyn). Coefficients are invariant to the Conley cutoff by construction; only standard errors and the first-stage Wald F differ.

Table OA.C.2: Alternative Inference: Wild Cluster Bootstrap (OLS) and Anderson–Rubin (IV)

	OLS (Province FE)	IV (No FE)
Model:	(1)	(2)
Variables		
Local Officials	0.0071**	0.0165***
	(0.0024)	(0.0044)
WCB 95% CI	[0.0003, 0.0130]	—
AR 95% CI	—	[0.0080, 0.0298]
p-value	0.043	0.003
Controls	Yes	Yes
FE	Province	—
Fit statistics		
Observations	1,132	1,132

Asymptotic clustered (district) SE in parentheses
Stars: col. (1) WCB p-value; col. (2) AR p-value. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Robustness check for Table 2. Column (1): wild cluster bootstrap (Cameron et al., 2008), Rademacher weights, $B = 9{,}999$ replications, 93 district clusters; p-value and 95% CI from the bootstrap distribution. Column (2): Anderson–Rubin test, which is robust to both few clusters and weak instruments ($F = 9.4$, $p = 0.003$). Point estimates are identical to Table 2.

Instrument Validity

To assess the sensitivity of the IV estimates to violations of the exclusion restriction, I implement the plausibly exogenous bounds of Conley et al. (2012). This approach asks: how large could the instrument's direct effect on the outcome be before the IV estimate loses significance? Calibrating the maximum direct effect at 50% of the reduced-form coefficient, I compute the union of 95% confidence intervals over a grid of direct-effect values. For the Baturyn distance instrument with district-clustered standard errors, the union CI is $[0.000, 0.036]$: the point estimate remains positive across the entire grid and the lower bound is essentially zero, indicating that the IV result is not sensitive to moderate violations of the exclusion restriction. The Anderson–Rubin test, which is robust to weak instruments, rejects the null of zero effect ($F = 9.4$, $p = 0.003$; 95% CI $[0.008, 0.030]$).

Sensitivity to Unobservable Confounding

I assess the sensitivity of the OLS estimates to omitted variable bias using the framework of Cinelli and Hazlett (2020). The robustness value (RV) for the main OLS specification is 12.1%, meaning that an unobserved confounder would need to explain at least 12.1% of the residual variation in *both* the treatment and the outcome to fully explain away the estimated effect. The RV for statistical significance at the 5% level is 6.7%. Benchmarking against observed covariates, the two strongest are Ukrainian population share ($R^2_{D\sim Z|X} = 20.1\%$) and merchants per 1,000 (12.7%). A confounder as strong as either covariate leaves the estimate significant ($\hat{\beta} = 0.0047$, $t = 2.6$ benchmarking against ethnicity; $\hat{\beta} = 0.0050$, $t = 2.8$ benchmarking against merchants). Both benchmarks are conservative because neither covariate independently predicts the outcome conditional on the treatment ($R^2_{Y\sim Z|D,X} = 0.007$ for ethnicity and 0.010 for merchants). Figure OA.C.1 marks each benchmark's partial R^2 at 1× and 2× its observed strength: at 1× both estimates remain significant at the 5% level, while at 2× both stay positive but are no longer significant. These results indicate that the OLS estimates are reasonably resilient to omitted variable concerns.

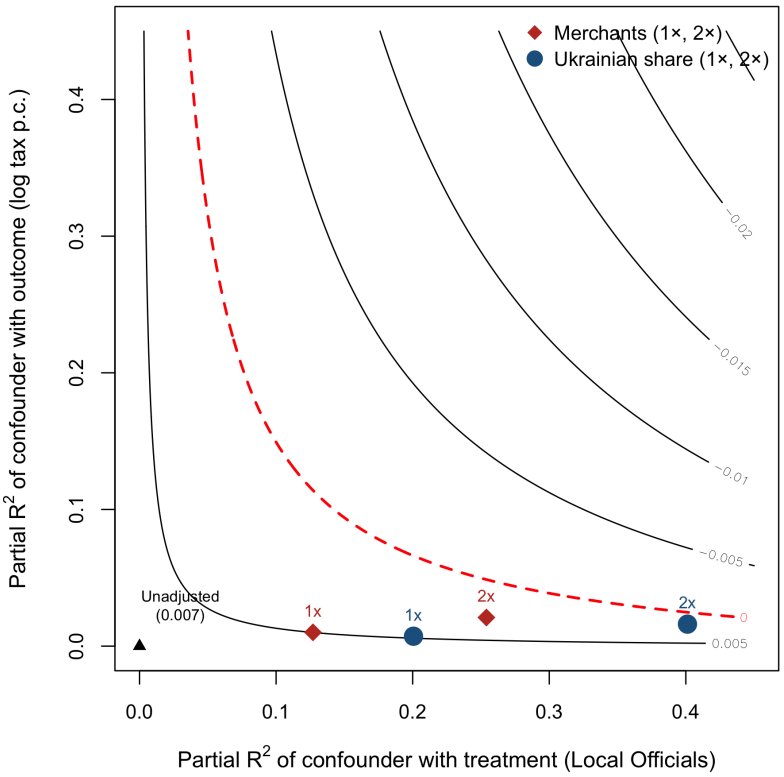

Figure OA.C.1: Sensitivity of the OLS estimate to unobserved confounding

Notes: Contour lines show how the estimated effect changes as a hypothetical confounder's partial R^2 with the treatment (x-axis) and outcome (y-axis) increases (Cinelli & Hazlett, 2020). The red dashed contour marks zero effect. Red diamonds (merchants per 1,000) and blue circles (Ukrainian population share) mark each covariate's partial R^2 at 1× and 2× its observed strength. Based on the OLS specification with province fixed effects ($N = 1{,}132$).

Capital City vs. Rural Officials

A natural concern is whether Ukrainian officials held real administrative authority. In the Russian Empire, executive and judicial power concentrated in provincial and district capital cities, where governors, courts, and police headquarters were located. If Ukrainians served only as rural clerks while Russians staffed the capitals, the treatment variable would overstate Ukrainian governance influence. The 1897 census records officials for every town in each district; I use officials in the district capital city to test this concern. Table OA.C.3 disaggregates the treatment variable. Ukrainian representation was lower in capital cities (mean 53%) than in rural areas, defined as the parts of each district outside any town (mean 62%), but both components predict contemporary tax revenue: in separate regressions with province fixed effects, the capital-city share yields a coefficient of 0.003 ($p < 0.10$) and the rural share yields 0.004 ($p < 0.05$). In the IV design, the Baturyn instrument predicts Ukrainian representation in capital cities ($F = 13.1$) more strongly than the district aggregate ($F = 11.8$), and the second stage is significant ($\hat{\beta} = 0.011$, $p < 0.01$). The district-level aggregate remains the preferred measure because it uses all available variation and avoids measurement error from splitting small counts.

Table OA.C.3: Ukrainian Official Representation by Administrative Level (1897)

	Aggregate	Capital City	Rural
Panel A: Descriptive Statistics (district level)			
Mean (%)	52.5	52.8	62.2
SD	20.6	23.4	16.5
Min	11.6	12.0	22.4
Max	95.7	98.9	93.8
Corr. with aggregate	1.00	0.95	0.71
N districts	93	93	93
Panel B: OLS with Province Fixed Effects			
Coefficient	0.0071^{***}	0.0033^{*}	0.0038^{**}
	(0.0024)	(0.0018)	(0.0017)
N	1,132	1,132	1,132
Panel C: IV (Baturyn instrument, no fixed effects)			
Coefficient	0.0165^{***}	0.0110^{***}	
	(0.0044)	(0.0034)	
First-stage F	11.8	13.1	
N	1,132	1,132	

Clustered (district) standard-errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Notes: Aggregate uses Ukr/Total (all ethnicities), matching the paper's main specification. Capital city and rural columns use Ukr/(Ukr+Rus) because total officials are not recorded separately for towns; the two measures correlate at $r = 0.95$ at the district level. Capital city is the town in each district with the most officials (the administrative seat in nearly all cases). Rural covers the rest of the district outside towns. Panel A: 93 districts in the estimation sample. Panel B: DV is log(local tax per capita, 2021). Province FE (11 provinces). Panel C: Instrument is log distance to Baturyn. First-stage F is the Wald statistic. Rural IV omitted due to weak first stage.

Notes

1. Specification: full geographic, soil, and historical controls plus region fixed effects; HC1 standard errors; $N = 93$. The LRCC-DVNL 2021 raster exhibits a distributional discontinuity inconsistent with the preceding series (the share of zero-valued pixels increases from 52% to 67% and mean luminosity drops by 31%) and is therefore excluded.

References

Arends, H., Brik, T., Herrmann, B., & Roesel, F. (2023). Decentralization and trust in government: Quasi-experimental evidence from Ukraine. *Journal of Comparative Economics*, *51*(4), 1356–1365. https://doi.org/10.1016/j.jce.2023.08.002

Buggle, J. C., & Nafziger, S. (2021). The Slow Road from Serfdom: Labor Coercion and Long-Run Development in the Former Russian Empire. *The Review of Economics and Statistics*, *103*(1), 1–17. https://doi.org/10.1162/rest_a_00857

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, *90*(3), 414–427. https://doi.org/10.1162/rest.90.3.414

Cinelli, C., & Hazlett, C. (2020). Making Sense of Sensitivity: Extending Omitted Variable Bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(1), 39–67. https://doi.org/10.1111/rssb.12348

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*(1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0

Conley, T. G., Hansen, C. B., & Rossi, P. E. (2012). Plausibly Exogenous. *Review of Economics and Statistics*, *94*(1), 260–272. https://doi.org/10.1162/REST_a_00139

Food and Agriculture Organization of the United Nations. (2003). Digital Soil Map of the World [Digital version 3.6, completed January 2003; originally published as the FAO/UNESCO Soil Map of the World at 1:5 000 000 (1974–1978); copyright FAO/UNESCO 1995].

Hajda, L., Kryzhanivsky, S., Yerofeyev, I., Zasenko, O., Makuch, A., & Stebelsky, I. (2025). The Cossacks of Ukraine. *Encyclopedia Britannica*.

Harus, A., & Nivyevskyi, O. (2020). In Unity There Is Strength: The Effect of the Decentralization Reform on Local Budgets in Ukraine. *VoxUkraine*.

Harvard Ukrainian Research Institute. (2024). MAPA: Digital atlas of ukraine — famine web map [Data retrieved 2026-03-27].

Hatsko, V., Brik, T., Tytiuk, S., Koval, A., Piddubnyi, I., Petrynka, K., Savisko, M., & Darkovich, A. (2025, April). KSE Local Data Hub. https://doi.org/10.5281/ZENODO.15267573

Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe version 4 [Hole-filled SRTM 90 m elevation database; 30 arc-second aggregate accessed via the geodata R package].

Kelso, N. V., & Patterson, T. (2010). Introducing Natural Earth data — naturalearthdata.com [10 m coastline accessed via the rnaturalearth R package]. *Geographia Technica*, (Special Issue), 82–89. https://www.naturalearthdata.com

Kessler, G. a. A. M. (2025, February). Electronic Repository of Russian Historical Statistics, 18th - 21st centuries, https://ristat.org/ Version I (2020): Russian Empire Historical GIS Maps (1897).

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems [HydroRIVERS v1.0; data via HydroSHEDS, https://www.hydrosheds.org]. *Hydrological Processes*, *27*(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Markevich, A., Naumenko, N., & Qian, N. (2025). The Causes of Ukrainian Famine Mortality, 1932–33. *Review of Economic Studies*, *92*(5), 3276–3305. https://doi.org/10.1093/restud/rdae091

Naumenko, N. (2026). Economic consequences of the 1933 Soviet famine. *Explorations in Economic History*, *100*, 101738. https://doi.org/10.1016/j.eeh.2025.101738

Rabinovych, M., Brik, T., Darkovich, A., Savisko, M., Hatsko, V., Tytiuk, S., & Piddubnyi, I. (2024). Explaining Ukraine's resilience to Russia's invasion: The role of local governance. *Governance*, *37*(4), 1121–1140. https://doi.org/10.1111/gove.12827

Reznik, O. M., Riadinska, V. O., Kuybida, V. S., Zubenko, H. V., & Zastrozhnikova, I. V. (2020). Regulatory and Organizational Aspects: Administrative-Territorial Reform in Ukraine. *International Journal of Management*, *11*(4), 576–584.

Tang, H., Zhong, Y., & Xia, H. (2025, April). Global nighttime light dataset from 1992 to 2022 with focus on low-light areas. https://doi.org/10.7910/DVN/15IKI5

Trojnitskij, N. (1903—1905). *Pervaya Vseobshchaya Perepis' Naseleniya Rossijskoj Imperii 1897 Goda. (Gubernskie Itogi) [Первая всеобщая перепись населения Российской Империи 1897 Года (Губернские Итоги)]* (T. 1–89). Statotdel TsSK press.

U-LEAD. (2025, September). Byudzhety Terytorialnykh Hromad Ukrayiny [Detsentralizatsiya v Ukrayini Бюджети Територіальних Громад України].

WorldPop & Bondarenko, M. (2020). Individual Countries 1km UN Adjusted Population Density (2000-2020). https://doi.org/10.5258/SOTON/WP00675

Yasnopolskij, N. (1913). *Statistiko-Ekonomicheskiye Ocherki Ekonomiki, Guberniy i Gorodov Rossii [Статистико-экономические очерки областей, губерний и городов России]* (Vol. 1). L.M. Fish.

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N.-E., ... Arino, O. (2021). ESA WorldCover 10 m 2020 v100 [Global land-cover map at 10 m resolution; "trees" class used for forest cover]. https://doi.org/10.5281/zenodo.5571936