

Control through Empowerment: Evidence from Nation-Building in Soviet Central Asia

Paul Castañeda Dower^a Andrei Markevich^b Ekaterina Zhuravskaya^c

^aUniversity of Wisconsin–Madison

^bNew Economic School

^cParis School of Economics (EHESS)

March 2023

Motivation: Colonization, development, and conflict

Does political empowerment of colonized groups exacerbate or mitigate conflict?

- Throughout colonial history, empires relied on a combination of coercion and cooptation
 - The question of whether to invest in local development sparked heated debates
 - It transcended economic and political considerations, bringing in the moral aspects of the “White Man’s Burden” (Rudyard Kipling, 1899)
- One key political-economy issue in this debate was whether political and cultural empowerment of colonies threatened the empire’s existence

We want to shed light on it using a unique historical experiment conducted by Bolsheviks in Central Asia

Motivation: From the Russian Empire to the USSR

The 1917 Revolution made the Soviet state an heir of the Russian Empire

- For economic, political, and military reasons, the Bolsheviks were determined to keep control over the Empire's colonies
 - Formally, Communist ideology condemned imperialism
 - Independence movements during the civil war showed that the Bolsheviks were unable to keep the centralized control
- The Bolsheviks turned empire building on its head by empowering colonial indigenous groups with autonomy
 - a union of autonomous nations rather than a unitary centralized state
 - Several republics within the USSR were comprised of groups that had well-defined preexisting national identity, such as Ukrainian Soviet Republic (1917)

Motivation: The Inception of Central Asian States

- The formation of the USSR also involved creating five Central Asian republics from scratch
- Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan
 - without pre-existing borders
 - without pre-existing national identities
 - clan, lineage, and religion were the primary identities
 - with few written languages
 - widespread illiteracy (e.g., Bartold 1927)

Soviet Experiment in Central Asia in the 1920s

Three components of the reform (launched in 1924–1925):

- ① Delimitation that created republic borders in Central Asia
- ② Political empowerment of the “titular nation” in each republic
 - Official recognition of their autonomy within the Soviet Union, de facto the establishment of nation states
 - Preferential access to resources and government / party jobs for the titular national group
- ③ Large-scale implementation of affirmative action policies toward the titular nation
 - One of the largest liquidation of illiteracy campaigns in history
 - Massive investments in schools together with the creation of written languages and promotion of local culture

Research question and preview of the main results

Aim:

- Shed light on the effect of the creation of Central Asian republics on:
 - ① the insurgency against the Soviet state (i.e., political conflict)
 - ② local inter-group conflict (i.e., ethnic conflict)
 - both important for the stability of the Soviet empire
 - ③ the nation building of Central Asian nations

Preview:

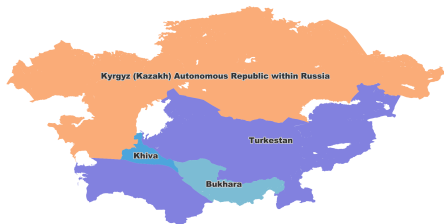
- Borders were drawn very carefully according to homelands of linguistic groups
- Reform was associated with a reduction of both ethnic and political conflicts relative to local areas where indigenous groups got local autonomy earlier
 - In contrast, conflict sparked in the region with unmet expectations for autonomy
- Changes in naming patterns and identities reported in censuses suggest that reform caused stronger identification with titular national groups
 - Schools were an important channel

Map of Central Asia Before the Reform



Map of Central Asia After the Reform





- **Before 1924 reform:** Two different national projects in Central Asia after the empire collapsed: (1) the Kazakh nation and (2) the Muslims of Turkestan" (Khalid 2021)
 - Pre-1924 North Kazakhstan (NK): was part of Russia not Central Asia
 - Colonized earlier; voted to join Russia as autonomous region before Oct 1917
 - 1920: Kyrgyz ASSR within Russia (Kazakhs were called Kyrgyz then)
 - "The national principle is completely alien to local... traditions" [of Turkestan] (Bartold 1925)
 - "[even young people could] not solve... an elementary question as what to call their people—Turk, Uzbek, or Muslim" (*Life of Nationalities*, Sep 1920)
- **After 1924 reform:** New borders, new ethnic identities everywhere, arguably with the exception of North Kazakhstan

Contribution to the Economic Literature

Rohner and Zhuravskaya (2023) (ed.) CEPR volume describing the state-of-the-art in the literature on nation building

- National self-identification

- Alesina & Spolaore 1997, 2003; Alesina et al. 2011; Michalopoulos & Papaioannou 2016; Bandyopadhyay & Green 2013; Bazzi et al. 2019
 - The opposite to the image of the “Artificial States”

- Nation-building policies

- Alesina & Reich 2015; Bandiera et al. 2019; Fouka 2019
 - Soviet Central Asian development policies were unprecedented in scope

- Population heterogeneity and conflict

- Easterly & Levine 1997; Alesina et al 2016; Bazzi & Gudgeon 2021; Lecce et al. 2022
 - Evidence that reform worked less well in more polarized places

► Historical Debate on Soviet Nationalities Policy

Several building blocks of the data: panel of grid cells

- ① Digital maps: Drawn from official paper maps
 - use grid cells of 0.1×0.1 degree resolution, about 40,000 [▶ grid cells](#)
 - need it to take into account multiple re-drawings of borders [▶ borders](#)
- ② Conflicts: Recently-declassified secret police (NKVD) reports, 1922 – 1934
 - 16 volumes of text describing the local situation on the ground within the union with special section devoted to Central Asia
 - Classify them manually and using machine text analysis (LDA) [▶ examples](#)
- ③ Names of Soviet citizens: WWII draftees (Pamyat Naroda) and victims of political repression (Memorial)
 - Last names proxy ethnicity; first names proxy ethnic self-identification
- ④ Population structure: Censuses and surveys
 - Pre- and Post-reform linguistic composition, 1922 and 1926
 - Post-reform ethnic composition, 1926 and 1939
- ⑤ Nation-building inputs: Soviet statistics
 - Various years during 1914 – 1940 (mostly in the 1930s)
 - Regional (more aggregated) panel data

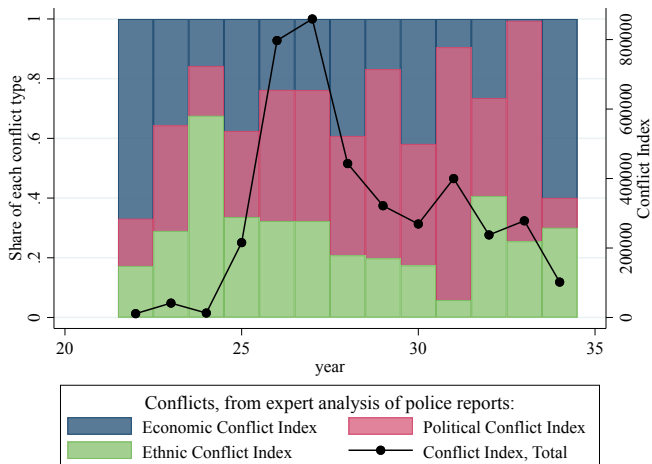
Manual classification of conflicts from secret police reports

- Broad definition: local tensions between groups of agents, individuals or the state
- Total of 5,430 entries in the conflict dataset in Central Asia
- Classified of all conflicts into:
 - **Political conflicts:** expressions of discontent with and riots (or insurgency) against the Soviet power or its local representatives
 - **Ethnic conflicts:** tensions and conflicts between different clans or groups
 - Other, for instance, economic, such as land disputes
- Geolocalized to the corresponding administrative unit and year mentioned in the report
- Aggregated conflict data into indices, dividing by the maximum number of mentions of any conflict over all grid cell \times year

	Mean	St. Dev.	Min	Max	N
Ethnic Conflict Index	0.017	0.043	0	0.532	532,974
Political Conflict Index	0.030	0.071	0	0.956	532,974

Conflict increased after the Reform, then subsided

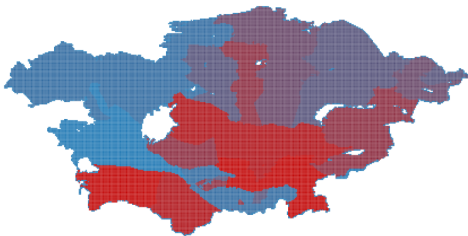
We want to see how it is distributed spatially



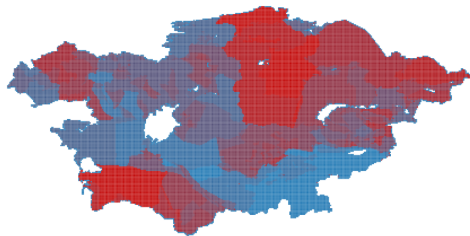
► Similar pattern with LDA-coded conflict

Spatial distribution of Political and Ethnic conflicts

Pre-reform:



Post-Reform:



Legend: blue – below mean, red – above mean, shades represent deciles

How to find a counterfactual?

- The reform defined the titular national group in newly created republics in the reform area; this changed the political status of every citizen: sorting them into titular or non-titular ethnicity
- In contrast, arguably, the reform did not change the status of residents of North Kazakhstan locally:
 - Kazakhs were the titular national group and others were not, already starting in 1920
- If one could abstract from expectations about which groups should be empowered, one could consider North Kazakhstan as a control group

Which groups should have received autonomy?

Were there any groups that hoped to get autonomy and did not? The unmet expectations could spark conflict

- We designed a local-search algorithm to construct an “ideal division” into republics allowing for their endogenous number
- Feed the machine with data on pre-reform local linguistic composition and apply the following criteria:
 - Group “like with like” and keep obtained clusters (republics) as distinct as possible
 - Specify a cost function that minimizes the number of unclassified cells per cluster (republic)
 - Subject to a minimum population size (400,000 people) and the presence of a city (of at least 10,000 people)
 - Constraints were stated by the Bolsheviks. They reflect the economies of scale à la Alesina & Spolaore 2003
- The solution is invariant to the starting point

▶ Algorithm details

Ideal partition: an outcome of the algorithm

6 states, not 5! Otherwise, close to the real partition, opposite of “Artificial states”



Ideal partition: an outcome of the algorithm

Real division, for comparison:



Ideal partition

Algorithm: grey spots are missing data, e.g., deserts, that the algorithm cannot assign



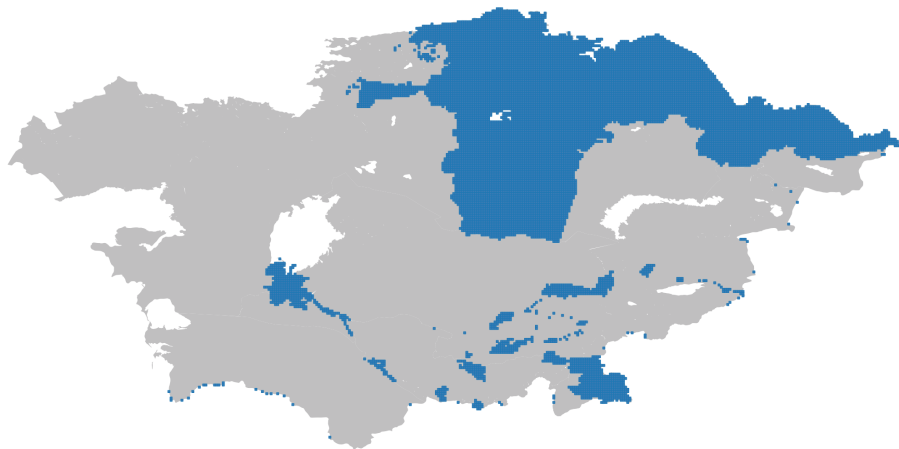
Ideal partition

If assign grey spots the same way as real:



Compliers (grey) and Non-compliers (blue)

Karakalpakiya and Gorny Badakshan got sub-autonomy, so could be treated as compliers



Empirical strategy, taking into account expectations

- The population of the “ideal” European republic (the non-complier area in North Kazakhstan) may have had unmet expectations for getting autonomy, as it is distinct from the rest of Kazakhstan
- In that case, this is not a good control group
- Thus, we consider areas that are the compliers in North Kazakhstan as the control group, instead

So, we estimate the simple difference-in-differences regressions with grid and year fixed effects with the main treatment being reform areas post 1924 and the control group being compliers in North Kazakhstan.

Decline in conflict relative to compliers in NK

European backlash

Dep. Var. =	Ethnic Conflict Index				Political Conflict Index			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Reform	-0.031** (0.012)	-0.019** (0.008)	-0.018** (0.008)	-0.021** (0.009)	-0.035** (0.016)	-0.019* (0.011)	-0.017 (0.011)	-0.024** (0.010)
Post × Control × Non-Complier		0.031*** (0.008)	0.031*** (0.008)			0.040*** (0.005)	0.040*** (0.005)	
Post × Reform × Non-Complier			-0.002 (0.002)				-0.018 (0.012)	
Post × Control × Share Europeans				0.040*** (0.011)				0.054*** (0.014)
Post × Reform × Share Europeans				0.009*** (0.003)				0.050 (0.041)
Cell & Year FEs, 1922 Rep. Trends	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03
R ²	0.20	0.22	0.22	0.21	0.26	0.27	0.27	0.27
Observations	532,974	531,349	531,349	532,974	532,974	531,349	531,349	532,974

- Large fall in conflict in the Reform area relative to North Kazakhstan, yet
- A much smaller decline in conflict relative to Compliers in North Kazakhstan (44% of SD for Economic and 27% of SD for Political Conflicts)
- Increase in conflict for noncompliers in NK: Unmet expectations for autonomy? Worsened discrimination of Europeans? ▶ Noncompliers in NK are Europeans

Acknowledgement of cost of the reform for Russians

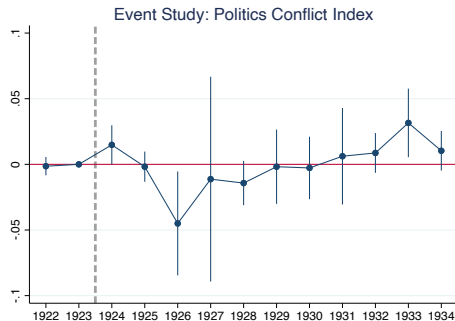
Nikolai Bukharin (1923):

“As the former Great Power nation [Russians], we should indulge the nationalist aspirations [of the non-Russians] and place ourselves in an unequal position, in the sense of making still greater concessions to the national current. Only by such a policy, when we place ourselves artificially in a position lower in comparisons to others, only by such a price can we purchase for ourselves the trust of formerly oppressed nations.”

(XII Congress (1923), 613 in Martin 2001, p. 17.)

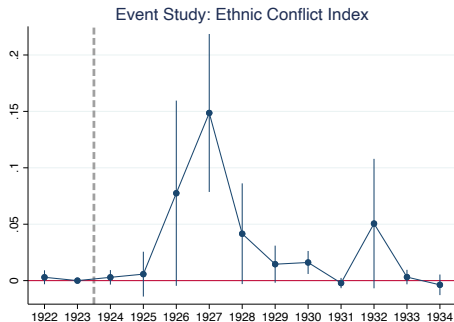
Event Studies: Conflicts in Reform Group

Relative to Compliers in North Kazakhstan



Event study: Conflict for non-Compliers in Control

Relative to Compliers in North Kazakhstan



- Adjustments to the reform: creation of local autonomous districts around 1928, particularly in noncomplier areas, which did appease ethnic tensions

► Details of Reform adjustments

► Results on Reform adjustments

Outcome: LDA classification of conflicts, 1922-1929

Dependent variable: each paragraph is attributed to a place and assigned a dummy for a topic if the probability of this topic >50%; then, aggregate all paragraphs that mention a particular place and divide by maximum over the sum of all topics.

Dep. Var. =	Topic 1: Political organized (1)	Topic 2: Political spontaneous (2)	Topic 3: Ethnic (3)	Topic 4: Economic I (4)	Topic 5: Economic II (5)
Post × Reform	-0.029*** (0.004)	0.009 (0.009)	-0.021*** (0.003)	-0.002 (0.002)	-0.002 (0.002)
Post × Control × Non-Complier	0.001 (0.001)	0.001 (0.002)	-0.003 (0.002)	0.001 (0.002)	-0.001 (0.001)
Cell & Year FEs; 1922 Rep. Trends	✓	✓	✓	✓	✓
Mean Dep. Var.	0.03	0.07	0.02	0.01	0.01
R ²	0.49	0.90	0.56	0.76	0.68
Observations	326,984	326,984	326,984	326,984	326,984

- Fall in organized political conflicts and ethnic conflicts after the reform
- Partial support for the evidence from hand-coded conflict outcome
- No evidence of the European backlash: probably, >50% is too crude measure

Heterogenous effects by local linguistic polarization

Dep. Var. =	Ethnic Conflict Index		Political Conflict Index	
	(1)	(2)	(3)	(4)
Post × Reform	-0.014* (0.008)	-0.014* (0.008)	-0.030*** (0.010)	-0.030*** (0.010)
Post × Control × Polarization	0.022*** (0.005)	0.025*** (0.005)	0.012*** (0.004)	0.012*** (0.004)
Post × Reform × Polarization	-0.004 (0.003)	-0.004 (0.003)	0.042** (0.020)	0.042** (0.020)
Post × Control × Non-Complier	0.016* (0.008)	0.072 (0.073)	0.033*** (0.006)	0.048 (0.052)
Post × Control × Polarization × Non-Complier		-0.071 (0.088)		-0.019 (0.062)
Cell & Year FEs, 1922 Rep. Trends	✓	✓	✓	✓
Mean Dep. Var.	0.02	0.02	0.03	0.03
R ²	0.22	0.22	0.27	0.27
Observations	531,349	531,349	531,349	531,349

- Polarized localities in the Reform – smaller decline or increase of political conflict, not ethnic conflict: Blame the Soviets for singling out the other group among equals?

► No heterogeneity by the share of TN

► Map: share of TN

► Heterogeneity by shares of nomads and TN

Nation-building outcomes and inputs

Two competing objectives:

(1) Creating nation-states and (2) Sovietization

Nation-building: Two competing objectives

Vladimir Lenin (1916):

The goal of socialism is not only to destroy the divisions of humankind into small states and to reduce the isolation of nations, not only in making nations closer to each other, but in their merger.

Joseph Stalin (1929):

It would be a mistake to think that the first stage of the global dictatorship of the proletariat would initiate the withering away of the concept of nations ... On the contrary, the first stage ... will be the stage of growth and flourishing of previously oppressed nations and national languages, the stage of establishing the equality of nations ...

In 1961, Nikita Khrushchev claimed that a new Soviet nation emerged (even though the ethnicities that comprised it are also preserved).

Nation-building outcomes and inputs data

- Short-run outcome:
 - Classify first names: Titular National, Muslim, Russian, or Soviet
 - All WWII draftees born 1920-1927 (from Pamyat Naroda)
 - Ethnic names: using Forebears.io Onograph API and Memorial
 - Soviet and Muslim names: using available lists
- Medium-run outcome:
 - Share of self-reported titular nationals
 - Ethnic definition: 1926, 1939 censuses (both after the reform)
 - Linguistic definition: 1922, 1926 censuses
- Inputs:
 - Educational inputs: schools, teachers, and pupils
 - Cultural inputs: theaters, museums, and cinemas
 - Regional-level panel in constant 1939 administrative borders
 - Unbalanced, for years: 1914, 1927, 1933, 1938, 1939, 1940, 1941

► Details: Classification of ethnic names

► Raw trends of the use of TN and Russian first names

Nation-building, short-run effects: First Names of Draftees

Diff-in-Diff: Central Asians (by last name) in Reform region vs. in NK and Russia's Belt

Dep. Var.:	Probability that First Name is:				
	Titular National		Muslim	Russian	Soviet
	This Republic (1)	Other CA Republic (2)			
Panel A: CA father in reform region vs. in North Kazakhstan complier region					
Post × Reform	0.004 (0.004)	0.001 (0.005)	−0.0001 (0.003)	−0.008*** (0.002)	0.001 (0.002)
Mean Dep. Var.	0.20	0.22	0.09	0.05	0.02
R ²	0.12	0.07	0.03	0.25	0.03
Observations	220,995	220,995	220,995	220,995	220,995
Panel B: CA father in reform region vs. CA or Russian father in Russian belt region					
Post × Reform	−0.006** (0.002)	0.002* (0.001)	0.003*** (0.001)	−0.014*** (0.002)	−0.017*** (0.001)
Mean Dep. Var.	0.55	0.04	0.02	0.56	0.05
R ²	0.14	0.26	0.13	0.22	0.02
Observations	1,278,167	1,278,167	1,278,167	1,278,167	1,278,167
District and Birth-Year Fixed Effects	✓	✓	✓	✓	✓
Father’s First Name Fixed Effects	✓	✓	✓	✓	✓
Father’s Last Name is Not TN	✓	✓	✓	✓	✓

Nation-building, short term effects: First Names of Draftees

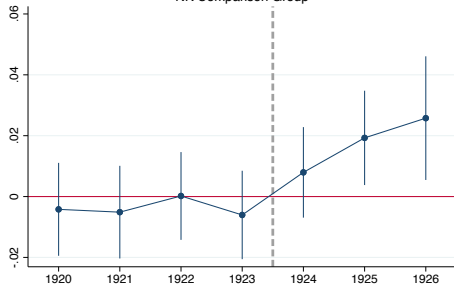
Heterogeneity by Father's Origin

Dep. Var.:	Probability that First Name is:				
	T titular National		Muslim	Russian	Soviet
	This Republic (1)	Other Republic (2)	(3)	(4)	(5)
Panel A: CA father in reform region vs. in North Kazakhstan Complier regions					
Post × Reform	-0.005 (0.006)	-0.001 (0.007)	-0.003 (0.005)	-0.002 (0.003)	0.003 (0.002)
Post × Reform × Father's Last Name is Not Titular National	0.019** (0.009)	0.001 (0.009)	0.007 (0.007)	-0.011** (0.004)	-0.006* (0.004)
Mean Dep. Var.	0.20	0.22	0.09	0.05	0.02
R ²	0.12	0.07	0.03	0.25	0.03
Observations	220,995	220,995	220,995	220,995	220,995
Panel B: CA father in reform region vs. CA or Russian father in Russian belt region					
Post × Reform	-0.012** (0.005)	-0.001 (0.002)	0.004*** (0.001)	-0.012*** (0.004)	-0.016*** (0.002)
Post × Reform × Father's Last Name is Not Titular National	0.013** (0.006)	0.008*** (0.002)	-0.002 (0.002)	0.008 (0.006)	0.003 (0.003)
Mean Dep. Var.	0.55	0.04	0.02	0.56	0.05
R ²	0.14	0.26	0.13	0.22	0.02
Observations	1,278,167	1,278,167	1,278,167	1,278,167	1,278,167
District and Birth-Year Fixed Effects	✓	✓	✓	✓	✓
All Double Interactions	✓	✓	✓	✓	✓
Father's First Name Fixed Effects	✓	✓	✓	✓	✓
Father's Last Name is Not TN	✓	✓	✓	✓	✓

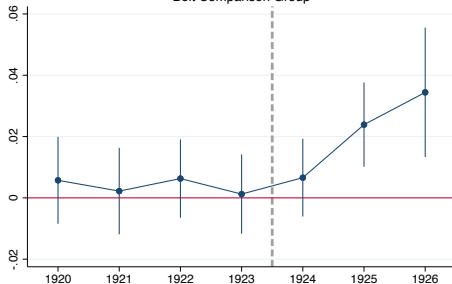
Event Studies: First names

Differential effect on the titular national first name for Central Asians with last names that are not titular national compared to those that are titular national

Event study: First name is Titular National
NK Comparison Group



Event study: First name is Titular National
Belt Comparison Group



Local Share of Titular Nationals as Medium-run Nation-Building Outcome

Cross-section of the first difference in the local share of titular nationals as outcome:

Dep. Var. =	Difference in the local share of titular nationals (among Non-Europeans)			
	between 1926 and 1939		between 1922 and 1926	
	Ethnic		Linguistic	
Definition of titular nationals:	(1)	(2)	(3)	(4)
Reform	0.084*** (0.029)	0.051 (0.053)	0.066 (0.070)	0.098* (0.056)
Control × Non-Complier	-0.062** (0.023)	-0.060** (0.024)	-0.021** (0.010)	-0.022** (0.010)
Reform × Share Titular Language < 0.33 in 1922		0.177** (0.085)		0.277** (0.131)
Reform × Share Titular Language > 0.66 in 1922		-0.010 (0.055)		-0.245*** (0.069)
Share Titular Language = 1 in 1922		-0.040 (0.030)		0.017 (0.072)
Mean Dep Var in Reform Area in Base Year	0.69	0.69	0.64	0.64
R ²	0.11	0.21	0.02	0.32
Observations	34,999	34,999	34,281	34,281

- Reform regions converged with North Kazakhstan in rates of identification with titular national group, particularly in areas where they were not a super-majority to start with
- Change of identity or migration?

► A Bird's Eye View of Population Changes: 1926 – 1939

Nation-building Inputs by Central Asian Republic

	No. of Schools		No. of Theaters		No. of Museums		No. of Cinemas	
	1914	1939	1914	1939	1914	1939	1914	1939
Uzbekistan SSR	160	4524	1	44	3	16	25	562
Turkmenistan SSR	58	1291	0	10	1	5	6	251
Tajikistan SSR	10	3834	0	22	0	2	0	122
Kyrgyzstan SSR	107	1754	0	19	0	3	1	187
Kazakhstan SSR	2011	8381	2	35	3	22	20	959

- INPRES reform in Indonesia (Duflo 2001) was modeled after Soviet development of Central Asia.
 - INPRES: In 1970s, roughly doubled number of primary schools (61,000 schools built in a country of over 100 million people).
 - Several aspects of the Soviet achievement more impressive: more schools per capita built, starting from a much lower base, and not just primary schools.

Nation-building inputs

Regional level panel:

Dep. Var. =	Educational Inputs (in logs)			Cultural Inputs		
	Number of Schools (1)	Number of Teachers (2)	Number of Pupils (3)	Number of Theaters (4)	Number of Museums (5)	Number of Cinema (6)
Post × Reform	1.85*** (0.40)	1.78*** (0.43)	2.38*** (0.70)	1.95** (0.81)	0.35 (0.34)	-10.71 (15.56)
Post × Control × Non-Complier	0.03 (0.05)	0.05 (0.04)	0.01 (0.04)	-0.16 (0.11)	0.13*** (0.04)	0.02 (2.80)
Regional Population (in logs)	0.40** (0.15)	-0.08 (0.17)	0.25 (0.19)	1.73* (0.98)	0.11 (0.25)	16.46 (14.85)
Region and Year Fixed Effects	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	5.70	6.97	10.28	2.73	1.00	47.74
R ²	0.82	0.87	0.76	0.57	0.35	0.67
Number of Regions	35	35	35	35	35	35
Observations	238	234	238	209	232	152

- Soviets invested a lot in education everywhere, but differentially more in reform regions.
- Cultural inputs also increased dramatically but not differentially so, except for museums in non-complier regions.
- Only one pre-reform cross-section, no way to test for pre-trends.

Are nation-building inputs important mediators?

Dep. Var. =	Local share of titular nationals among non-Europeans, 1926 and 1939						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Reform \times Year=1939	0.072** (0.032)	0.028 (0.039)	0.052 (0.036)	0.044 (0.036)	0.104** (0.040)	0.080** (0.036)	0.083** (0.037)
Control \times Non-Complier \times Year=1939	-0.122** (0.058)	-0.138** (0.057)	-0.134** (0.057)	-0.129** (0.056)	-0.128** (0.056)	-0.117* (0.057)	-0.125** (0.054)
No. of Schools (in logs)		0.046*** (0.017)					
No. of Teachers (in logs)			0.034** (0.015)				
No. of Pupils (in logs)				0.036*** (0.011)			
No. of Theaters					-0.010 (0.008)		
No. of Museums						-0.025* (0.014)	
No. of Cinemas							-0.001*** (0.000)
Regional population (in logs)	0.106** (0.041)	0.030 (0.036)	0.022 (0.037)	0.020 (0.036)	0.119*** (0.036)	0.122*** (0.033)	0.110*** (0.026)
Cell & Year FEs	✓	✓	✓	✓	✓	✓	✓
Mean Share of dep. var in Reform Area in 1926	0.69	0.69	0.69	0.69	0.69	0.69	0.69
R ²	0.31	0.32	0.32	0.32	0.31	0.32	0.34
Number of Clusters	31	31	31	31	31	31	31
Observations	70,350	67,865	67,865	67,865	68,445	68,445	68,445

- Education mediates the effect of the reform, while cultural inputs do not
- Education increased human capital and promoted local identity (in addition to Soviet one)
- Cinemas promoted Soviet identity, as movies were produced in the center (?)

Conclusions

- The national delimitation was done very carefully taking into account the linguistic composition of local indigenous population
- The creation of Central Asian republics reduced ethnic and political conflicts relative to Kazakh-dominated areas in North Kazakhstan
- The Central Asian population that was not titular national started to identify as titular national
- The share of titular nationals increased over time and differentially more so due to the reform and due to massive nation-building investment, such as school construction
- The results do suggest that the reform paid off for the Soviets (despite the European revolt in North Kazakhstan) as Soviets managed to keep control for 70 years

APPENDIX

Examples of Conflict Data

Volume 1, 1922, Report 3, Turkestan

Enver Pasha remains to be a central figure uniting those who are against Soviet policy. In military terms, all his efforts, now, aim to construct a large armed calvary and regular army. Enver mobilized Bukharan Basmachi, whom he gradually armed with the help of Afghanistan, to the area around the city of Bajsum in the eastern part of Bukhara.

- 1 Manual classification: “Crime,” “Central Authority,” “Open riot” → Political conflict
- 2 Machine-learning classification: Classification probabilities heavily concentrated on one latent “topic” → LDA Topic – Political conflict, organized

▶ back

Examples of Conflict Data

Volume 1, 1923, Report 3, Turkestan

In Poltoratsky subdistrict in Krasnovodsky district: The fighting between the Kyrgyz (Kazakh) and the Yomuds (Turkmen tribe) continues, it is becoming serious.

- 1 Manual classification: “Cross-ethnic,” “Ethnic discrimination” → Ethnic conflict
- 2 Machine-learning classification: Classification probabilities heavily concentrated on one latent “topic” → LDA Topic – Ethnic conflict

▶ back

Examples of Conflict Data

Volume 5, 1927, Report 2, Kazakhstan

52. Dzhetysujskaya province. Russian peasantry of the Ittifak volost in Dzarkent district expresses discontent with the slowdown in land management. There are disputed plots of land, which are simultaneously claimed by both Russians and Kyrgyz. Discontent intensified after the provocative speech of the UZU (district land management) secretary, who pointed to the “hopelessness of the expectations of the Russians.”

- ① Manual classification: “Cross-ethnic,” “Land reform implementation,” “Local official misbehavior,” “Land resources” → Ethnic and Economic conflict
- ② Machine-learning classification: Classification probabilities spread relatively equally across three latent “topics” → Unclassified

Data: Ethnic composition, pre-reform

To cover all of the Central Asia, collected data from different sources:

- 1920 Soviet population census with sub-district level data
 - Amu-Darya, Fergana, Samarkand, Syr-Darya provinces
- 1924 district-level survey for Bukhara and Khiva
- 1913 settlement-level data from a governor report for Semirechye province aggregated by sub-districts (+ 1921-22 Economic council reports)
- 1920-21 survey settlement-level data for a half of Kazakhstan aggregated by sub-districts
 - Bukeevsk, Aktyubinsk, Kustanaj, Turgaj provinces and Adaevsk district
- 1920 census district-level data for the rest of Kazakhstan
 - Aktyubinsk, Ural and Semipalatinsk provinces
- 1920 census district level-data for Turkmenistan

Empirical Strategy: Difference-in-differences

- Policy experiment

- Group treated with the reform: Turkestan, Khiva, Bukhara
- Naive Control group: North Kazakhstan
 - Take difference in the differences in pre/post reform outcomes between treatment and control groups
 - Note: The first post-reform year is 1924 (two pre-reform observations)

- Identification

- Assumption 1: In the absence of the reform, treatment and control would have had parallel trends
 - Historically-determined border between treatment and control
 - Endogenous timing?
- Assumption 2: Reform did not affect the Control group
 - Unmet expectations of a large non-Kazakh group, i.e., Europeans?

Basic Econometric Specification: Diff-in-Diff

$$Conflict_Index_{gt} = \gamma Post_t \times Reform_g + X_{gt}\beta + \alpha_g + \phi_t + \epsilon_{gt} \quad (1)$$

- *Reform* is the National Delimitation Reform Area (one in all grid cells in Turkestan, Bukhara, and Khiva; zero in North Kazakhstan)
- *Post* the onset of the reform, 1924
- *Conflict_Index* stands for indices of ethnic and political conflict in a grid cell g at time t .
- γ is the effect of interest.
- α_g and ϕ_t are grid and year fixed effects.
- X is set of controls
 - 1922 Republic linear trends
 - In baseline, only a dummy for 1933 famine in Kazakhstan (killed about a quarter of the pre-famine population)
- Standard errors clustered at the administrative district level

Progress in Education: Sweden vs. Uzbekistan

Educational outcomes of two countries with six million citizens

	Literacy Rate in 1914	University Attendance in 1938	Secondary School Attendance in 1938	Primary School Attendance in 1938
Sweden	99.7%	10,000	60,000	569,000
Uzbekistan	1.0%	17,500	179,000	916,000

Source: *The Soviet Far East and Central Asia*, Mandel, 1944

Implementation of National Delimitation in Central Asia (1924-1925)

- January 1924: Orgbureau of the Central Committee of the Communist Party discussed the delimitation that gave the start of the reform
- February - March 1924: Approved by Communist Parties of Bukhara, Khiva and Turkestan
 - Bukhara was split b/w Uzbekistan, Tadjikistan, Turkmenistan; Khiva was split b/w Uzbekistan and Turkmenistan
- June 1924: Approved by the Politburo of the Communist Party in Moscow
- Territory commission in Tashkent with national sub-commission on the ground in charge of border drawing
 - Moscow intervened only once and resolved the belonging of Tashkent
- October 1924: official announcement
- May 1925: Approved by the 3rd Soviet Congress in Moscow

► [Back to main slides](#)

Central Asia after the Russian Revolution

- Some discussion and preparations for national delimitation reform back to 1918
 - 1918: Turkestan Autonomous Republic (multi-ethnic)
 - 1918: Uzbek and Kazakh became official languages of Turkestan in addition to Russian; in 1919 Turkmen language was added
 - 1919: separate national departments for Uzbeks, Tadjiks, Turkmen, Kazakhs, Tatars, Ukrainians and Armenians were established within Turkomnats (Ministry of Nationalities of Turkestan) etc.
- 1920: Bolsheviks took control over Bukhara and Khiva (that used to be outside of the Russian empire)
- 1920: Kyrgyz (Kazakh) Autonomous Republic in North Kazakhstan
- By late 1923 Basmachi (anti-Soviet national movements) were largely destroyed
- Late 1923 - early 1924: national oblasts in Khiva and Bukhara

Official “Culturally-backward” Groups

- 97 nationalities of the Soviet Union, including all (non-European origin) Central Asian nationalities, were considered culturally backward because they met the following categories:
 - ① Low level of literacy
 - ② Low school enrollment
 - ③ Absence of written script with single literary language
 - ④ Presence of social vestiges: oppression of women, religious fanaticism, nomadism, racial hostility, clan vengeance
 - ⑤ Lack of specialists (national cadres of Soviet construction)
- The Bolsheviks claimed that national delimitation would promote building of Soviet nations in Central Asia and contribute to their development

Machine-learning Classification of Conflicts (LDA)

- Use all paragraphs that mention conflicts in Central Asia in all volumes of top secret reports before 1929
 - In 1929, the structure of the report changed
 - Use whole paragraphs in which text on Central Asia appears
- Perform Latent Dirichlet Allocation (LDA)
 - Unigrams, bigrams and trigrams
 - Summarize each “topic” using top 20 factor loadings
- Assign each entry, as before, to the corresponding administrative and time unit on the map (sub-district, district)
- Assign a topic for each entry if its probability is $>50\%$
- Sum up over entries and topics by grid cell and year, normalize by the largest sum of all topics in any g and t

LDA: 5 Topics

Topic 1: Political: Anti-Soviet rebellion, organized movements

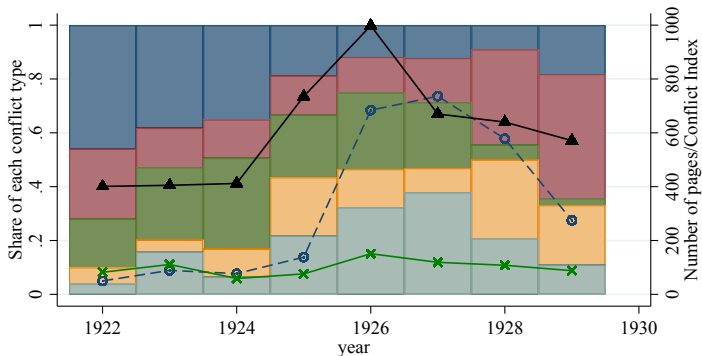
Topic 2: Political: Anti-Soviet protests, class conflicts, spontaneous

Topic 3: Ethnic conflicts

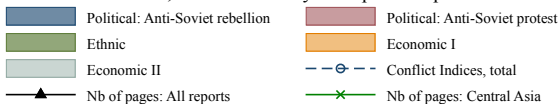
Topic 4: Economic conflict, rural: distribution of cattle and grain

Topic 5: Economic conflict, poverty and inequality

Composition of LDA topics by year



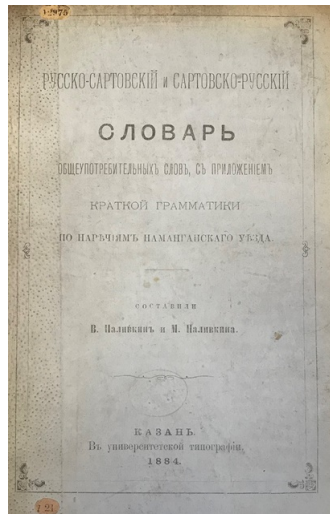
Conflicts, from LDA analysis of police reports:



[▶ back to main slides](#)

Ethnicities to Nations

- Language identification as a criteria for national republics
- But not always
 - Sarts became Uzbeks (same for Kurama)
 - Uyghurs, Kipchaks, Dungans and other small groups (in contrast to Karakalpaks) did not get their national autonomies
 - Kazakhs were first set up under the name of Kyrgyz, and Kyrgyz under the name of Kara-Kyrgyz



► Back to main slides

Entries by year and geolocation

Year	Number of entries	Share of entries geo-specified	
		at least as precise as district level	at least as precise as subdistrict level
1922	101	0.27	0.22
1923	166	0.34	0.22
1924	163	0.26	0.15
1925	261	0.78	0.69
1926	824	0.86	0.78
1927	878	0.85	0.71
1928	823	0.74	0.48
1929	774	0.80	0.47
1930	303	0.84	0.64
1931	538	0.83	0.78
1932	198	0.81	0.70
1933	176	0.68	0.60
1934	223	0.78	0.70
Total	5430	0.77	0.61

The reform and the precision of conflict information

Dep. Var. =	Precision Index All Conflict		Precision Index Ethnic Conflict		Precision Index Political Conflict	
	All levels	Subdistrict/ Settlement	All levels	Subdistrict/ Settlement	All levels	Subdistrict/ Settlement
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Reform	0.049 (0.067)	-0.066 (0.138)	0.014 (0.051)	-0.031 (0.137)	0.118** (0.049)	-0.136 (0.137)
Post × Share Titular National (bottom)	0.042** (0.017)	0.137* (0.071)	0.106** (0.041)	0.092 (0.084)	0.023 (0.029)	0.021 (0.058)
Post × Control × Non-Complier	-0.160* (0.094)	0.326** (0.140)	-0.027 (0.033)	0.132** (0.061)	-0.034 (0.046)	0.248*** (0.058)
1922 Rep. Linear Trends	✓	✓	✓	✓	✓	✓
Interactions with No Conflict Dummy	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.23	0.22	0.21	0.09	0.20	0.013
R ²	0.40	0.50	0.44	0.65	0.50	0.61
Number of Clusters	41	41	41	41	41	41
Observations	531,349	531,349	531,349	531,349	531,349	531,349

- Precision of geolocation of conflicts is going up, primarily to subdistrict and district levels in reform areas.
- Precision at the settlement level requires more titular nationals.
- Even columns show that we may be underestimating the effect of the reform.

Summary Statistics

	Mean	St. Dev.	Min	Max	N
Ethnic Conflict Index	0.017	0.043	0	0.532	532974
Political Conflict Index	0.030	0.071	0	0.956	532974
Reform Group	0.365	0.482	0	1	40998
Share Titular National	0.738	0.298	0	1	38409
Complier	0.717	0.451	0	1	40873
Adjustment Group	0.060	0.238	0	1	40998
Share Europeans	0.188	0.260	0	1	40998
Polarization Index	0.377	0.370	0	1	40998
Share Nomads	0.718	0.337	0	1	35448
Share Urban Pop.	0.073	0.119	0	1	29933
Share Illiterate Pop.	0.942	0.050	0.554	1	7697

Adjustments of Reform in Central Asia (1/3)

- Adjustments of the reform started right after the reform, but the major ones occurred in 1928
- The major adjustment of 1928:
 - National districts for minorities in national republics if they were local majorities: Russians, Ukrainians, Uzbeks, Kazakhs, Uyghurs etc.
- Other adjustments:
 - February 1926: Kara-Kyrgyz Autonomous Oblast was updated to Autonomous Republic and changed the name to Kyrgyz
 - October 1929: Tadjik Autonomous Republic was updated to Union Republic and got Khodzhent region
 - December 1936 (beyond of the period we have data for): Kyrgyz and Kazakhstan Autonomous Republic were updated to Union Republics

[▶ Results on Reform adjustments](#)[▶ Back](#)

Local Autonomy within Autonomous Nations: Who obtained adjustment? (1/2)

- Adjustments to the reform at lower-level administrative units
- Administrative districts received autonomy in 1928
 - Titular national districts for Russians and Ukrainians
 - but also Uyghurs, Karakalpaks, as well as Turkmens, Kyrgyzs, Kazakhs, Uzbeks in other republics

Cross-section of grid cells:

Dep. Var. =	Obtained Adjustment	
	(1)	(2)
Reform Group	0.069 (0.067)	0.091 (0.074)
Control \times Non-Complier	0.093** (0.039)	
Reform \times Non-Complier	0.025 (0.066)	
Control \times Share Europeans		0.161*** (0.058)
Reform \times Share Europeans		-0.245 (0.237)
Mean Dep. Var.	0.06	0.06
R ²	0.03	0.04
Observations	40,873	40,873

Experimentation: adjustments decreased ethnic conflict (2/2)

Dep. Var. =	Ethnic Conflict Index		Political Conflict Index	
	(1)	(2)	(3)	(4)
Post × Reform	-0.018** (0.008)	-0.017** (0.008)	-0.018* (0.010)	-0.016 (0.011)
Post × Control × Non-Complier	0.032*** (0.008)	0.034*** (0.009)	0.041*** (0.005)	0.042*** (0.005)
Post-1928 × Adjustment	-0.017 (0.015)	0.006 (0.008)	-0.011 (0.007)	-0.002 (0.005)
Post × Reform × Non-Complier		-0.002 (0.002)		-0.018 (0.012)
Post-1928 × Adjustment × Control × Non-Complier		-0.052*** (0.018)		-0.021 (0.016)
Post-1928 × Adjustment × Reform × Non-Complier		0.001 (0.007)		-0.003 (0.008)
Cell and Year FEs; Rep. Trends	✓	✓	✓	✓
Mean Dep. Var.	0.02	0.02	0.03	0.03
R ²	0.22	0.22	0.27	0.27
Observations	531,349	531,349	531,349	531,349

- No average effect
- The adjustments appeased the ethnic tensions in areas w/ European presence (also negative, but insignificant for political conflicts) [▶ Back](#) [▶ Details of Reform adjustments](#)

Distribution of Ethnic Groups in Real and Ideal Nations

	Share of titular nationals		Share of largest minority	
	Real	Ideal	Real	Ideal
Kyrgyzstan	0.578	0.719	0.286	0.173
Tajikistan	0.590	0.779	0.297	0.080
Turkmenistan	0.756	0.763	0.079	0.083
Uzbekistan	0.776	0.712	0.113	0.156
Kazakhstan	0.760	0.897	0.143	0.056
European republic	—	0.494	—	0.466

- All “ideal nations” except Uzbekistan are more homogenous than the real ones
- The European Republic is very polarized with 2 groups: Europeans and Kazakhs
 - The input to the algorithm did not aggregate different Europeans to the same group, the European Republic emerged as the algorithm’s output

Algorithm for the Ideal Partition

- 1 Randomly select a grid cell to start
- 2 Local search: join grid cell with a neighboring grid cell if distributions of linguistic groups are “close enough” to each other
 - within an ϵ neighborhood using Euclidean distance over all population groups
- 3 Form cluster when above search rejects all neighboring cells
- 4 Go back to Step 1 and repeat until no more clusters are formed
- 5 Optimally select what “close enough” means by choosing ϵ to minimize the number of unclassified cells per number of clusters
- 6 Merge cluster with the most similar neighbor if population size is too small or does not contain a city

► [back to the main slides](#)

Indices of Political and Economic Conflict

by year, covering 1922–1934, across 40,873 grid cells

- The conflict indices aggregate each entry about conflict for each grid cell g and year t

$$Conflict_Index_{gt}^{Type} = \sum_{i_v} \mathbf{1}(Type_{i_v}) \times \mathbf{1}(Location_{i_v}) \quad (2)$$

- i_v is the identifier for a particular entry i in volume v
- $Type_{i_v}$ indicates mention of an ethnic or political conflict in the entry
- $Location_{i_v}$ indicates entry i_v associated with g and t
- Information on conflicts has different geographic precision:
 - settlement, subdistrict, district; we cluster by district
 - some entries talk about provinces and republics, we ignore these entries in the baseline analysis
- Normalize by the maximum value of the sum of Political and Ethnic Conflict Indices across all g and t :
 - A value of one indicates the highest intensity of reporting in g and t
 - A value of zero indicates no reporting about conflict in g and t

	Mean	St. Dev.	Min	Max	N
Ethnic Conflict Index	0.017	0.043	0	0.532	532,974
Political Conflict Index	0.030	0.071	0	0.956	532,974

Russia's Colonization of Central Asia

Was gradually expanding:

- Key timeline difference is between the incorporation of Steppe region (North Kazakhstan) and Turkestan (Central Asia South of North Kazakhstan)
 - North Kazakhstan became part of the empire in the 18th century (by 1920s: over one century in the empire)
 - Turkestan was colonized in the second half of the 19th century (by 1920s: only 1/2 century in the empire)
 - Imperial definition of Central Asia did not include North Kazakhstan
 - It was part of Russia
 - More intense European migration (e.g., Russians, Ukrainians, and Germans), especially after the Stolypin reform (1906)
 - There is also difference in geography (steppe vs mountains and deserts)
- Within Turkestan, the empire had uniform policies
 - (Bukhara and Khiva remained independent until 1920, but were under protectorate)

The “National Question” in the early Soviet Union

- Class was the key political concept for the Bolsheviks
 - Initially, they viewed the ethnic divisions as the colonial instrument of control
 - Lenin called the Russian empire the “prison of nations” (1914)
 - nearly one half of the population of the empire was non-Russian
- As Soviet stated objectives were to undo all types of oppression, including colonial, initially they deemphasized the importance of ethnic divisions
- The rise of national movements after the 1917 revolution and during the Civil war forced the Bolsheviks to consider the “national question” seriously and gradually build a nationalities policy
 - In 1917 Stalin was appointed the *Commissar of Nationalities*

Conflicts from Secret Police Reports

- The Soviet secret police (VChK-OGPU-NKVD-MVD-KGB) closely monitored internal situation in the USSR
- The secret police regularly prepared top secret reports on the situation in the country for the highest Soviet authorities (texts only, no tables)
 - They were recently declassified and historians published 16 volumes containing *all* reports sent to Stalin in 1922-1934
 - Daily and monthly reports before 1929, ad hoc thematic reports after
 - Special section on “national regions,” including a subsection on Central Asia in 1922-1929
 - The procedure of gathering and aggregating information was highly centralized
 - There could be a possible selection problem in reporting, but (due to centralization) mostly over time

Manual Classification of Conflicts

- We manually created a dataset on all mentions of conflicts in Central Asia in the sixteen volumes of top secret reports
 - A broad definition of a conflict: local tensions between groups of agents (individuals or the state)
- Total of 5,430 entries in the conflict dataset
- Created classification of all conflicts along various dimensions with a special focus on:
 - Political conflicts (expressions of discontent with and riots or insurgency against the Soviet power, or its local representatives)
 - Ethnic conflicts (tensions and conflicts between different ethnic groups and clans)
- Geolocalized, assigning each conflict to the corresponding administrative unit and time (so far, years), mentioned in the report

Soviet Nationalities Policy

- Empower minorities that are local majorities via the establishment of autonomous republics
 - E.g., 1917 Ukrainian Soviet Republic, 1919 Belorussian Soviet Republic, 1920 Kazakh Autonomous Soviet Republic within Russia (on the territory of North Kazakhstan)
 - Early Soviet definition Central Asia also did not include North Kazakhstan (just as during Imperial period)
- After XII Party Congress in April 1923, indigenization (*Korenizatsiya*) policy was implemented in full scale in these republics
 - Promotion of local languages, culture, education in the local language
 - Local administration was shifting from Russian to local cadres and languages
- National delimitation of Central Asia (Turkestan, Bukhara, and Khiva) in 1924-1925 was the continuation of this Soviet Nationalities policy
- But as Central Asia did not have national borders, they needed to be created

Central Asia on the eve of the reform

- Lack of ethnic identity (Bartold 1927)
 - At best, lineage/clan and muslims identities
 - Example: Russian Empire did not make distinction between Kazakhs and Kyrgyz at all, Soviets called Kazakhs Kyrgyz, while Kyrgyz were called Kara-Kyrgyz. The North Kazakhstan before the reform was called Autonomous Kyrgyz Republic.
- Major languages
 - Turkic: Kazakh, Kyrgyz, Turkmen, Uzbek, Karakalpak
 - Iranian: Tadjik
- Uzbeks and Tadjiks were the most settled people
 - Ancient cities established during Alexander the Great (4th century BC)
- All the others traditionally were nomads
- Fergana Valley (with Tashkent and other cities) was the most developed region but also the most diverse (Uzbeks, Tadjiks, Kazakhs and Kyrgyzs)

National Delimitation in Central Asia (1924-1925)

- Idea: Transform ethnicities into Nations: Language identification as the main criterion for national republics ▶ exceptions
- January 1924: the official start of the reform: Decision of the Orgbureau of the Central Committee of the Communist Party
 - Some discussion of national delimitation dates back to 1918
 - ▶ Details on what preceded implementation
- Reform implemented during 1924 and the first half of 1925
 - Uzbekistan and Turkmenistan Union Republics
 - Tadjikistan Autonomous Republic (within Uzbekistan) and Kara-Kyrgyz Autonomous Oblast (within Russia)
 - South Kazakhstan became part of Kazakh Autonomous Republic
 - Karakalpakstan Autonomous Oblast (within Kazakhstan)
 - ▶ Details of implementation

Documenting administrative reforms

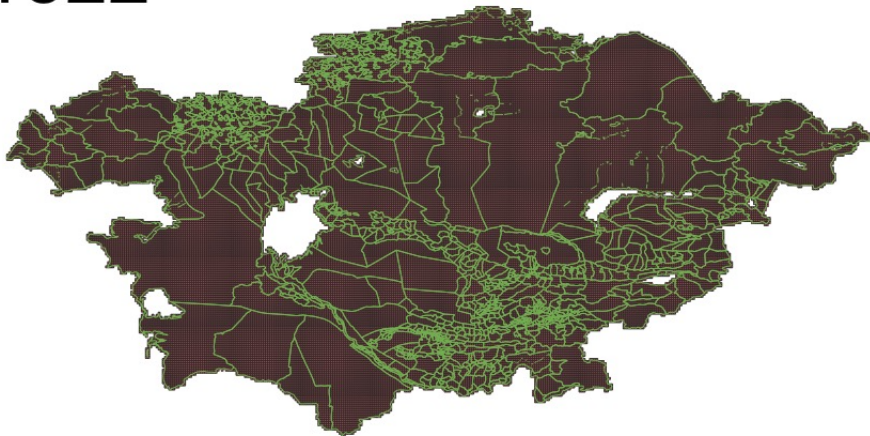
To create the dataset, need comparable over time spatial units

Stumbled on multiple administrative reforms, use grid cells of 0.1×0.1 degree resolution as a spatial unit of analysis

- To document border changes, we digitized many republic and province-level historical paper maps (including one that exists in single copy in the archives)
- Result: three digitized maps
 - 1922 pre-reform division
 - Match our ethnicity data with 1922 administrative borders
 - 1926 post-reform division
 - 1928 division post reform adjustments
- Account for transfers of low-level administrative units between higher-level administrative units during the periods between the reference-map years

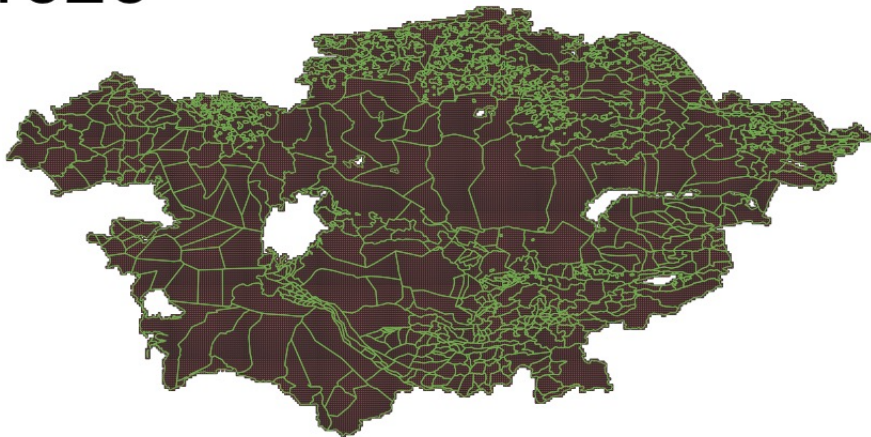
Border Changes: 1922 pre-reform division

1922

[▶ back](#)

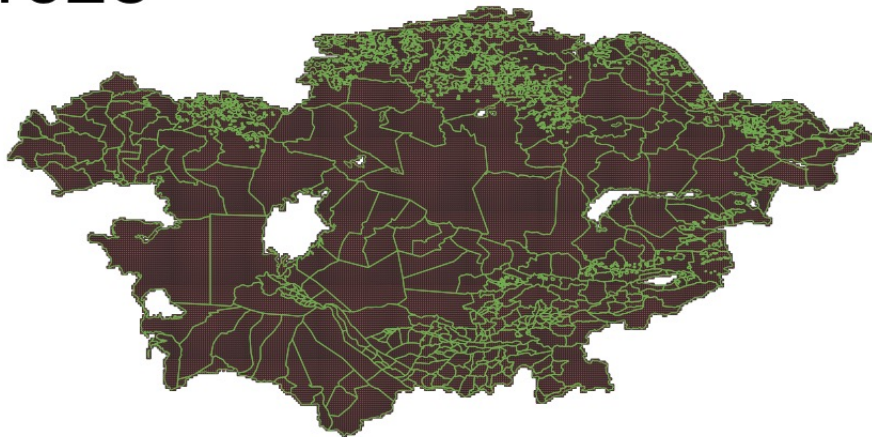
Border Changes: 1926 post-reform division

1926

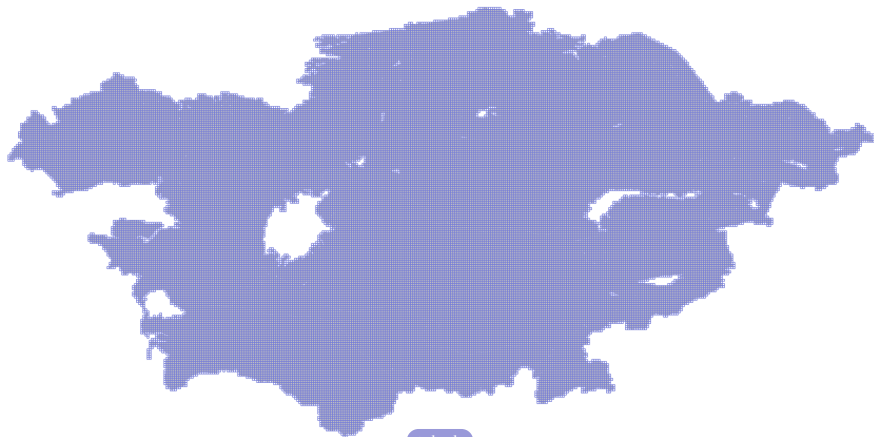
[▶ back](#)

Border Changes: 1928 post-reform adjustments

1928

[▶ back](#)

Grid cells

[▶ back](#)

Naive Diff-in-Diff estimates of the effect of the reform

Dep. Var. =	Ethnic Conflict Index		Political Conflict Index	
	(1)	(2)	(3)	(4)
Post \times Reform	-0.021*** (0.006)	-0.031** (0.012)	-0.016* (0.009)	-0.035** (0.016)
Cell and Year FEs	✓	✓	✓	✓
1922 Rep. Linear Trends		✓		✓
Mean Dep. Var.	0.02	0.02	0.03	0.03
R ²	0.20	0.20	0.26	0.26
Observations	532,974	532,974	532,974	532,974

- Decrease of both types of conflict relative to the Naive Control Group:
 - 70% of SD for ethnic conflict and 49% of SD for political conflict

Non-Compliers and Europeans

Cross-section of grid cells:

Dep. Var. =	Share Europeans	
	(1)	(2)
Reform Group	-0.027 (0.026)	-0.107** (0.052)
Control × Non-Complier	0.421*** (0.051)	0.201*** (0.060)
Reform × Non-Complier	-0.020 (0.025)	-0.066 (0.061)
Additional control variables	✓	
Mean Dep. Var.	0.19	0.19
R ²	0.53	0.74
Observations	40,873	40,873

Notes: Additional controls: pre-reform shares of urban population, nomads, and illiterate.

- Non-compliers in the Control group is the region with high European presence
- Europeans are few in the “Reform area”
- Non-compliers in Reform group are not European districts

Some Elements of Background

- Only Turkestan, Khiva, and Bukhara were considered as Central Asia before the reform, while North Kazakhstan was considered part of Russia.
 - It was colonized earlier, in the 18th century, i.e., about 50 years before Turkestan.
- Indigenous population of North Kazakhstan was given a local autonomy by the Soviets within Russia already in 1920, when they formed Kyrgyz Autonomous Soviet Republic within Russia
 - with Kazakhs as a titular nation, whom they called Kyrgyz.
 - Russian Empire did not make distinction between Kazakhs and Kyrgyz at all, while Soviets, in the early 1920s called Kazakhs Kyrgyz and Kyrgyz Kara-Kyrgyz.
- The borders of National Republics in the Central Asia (south of North Kazakhstan) needed to be drawn, unlike for other for other autonomous republics, this is why the formation of Central Asian republics was delayed until 1924–1925.
 - The borders were created taking into account local languages of indigenous populations.

Historical Debate on Soviet Nationalities Policy

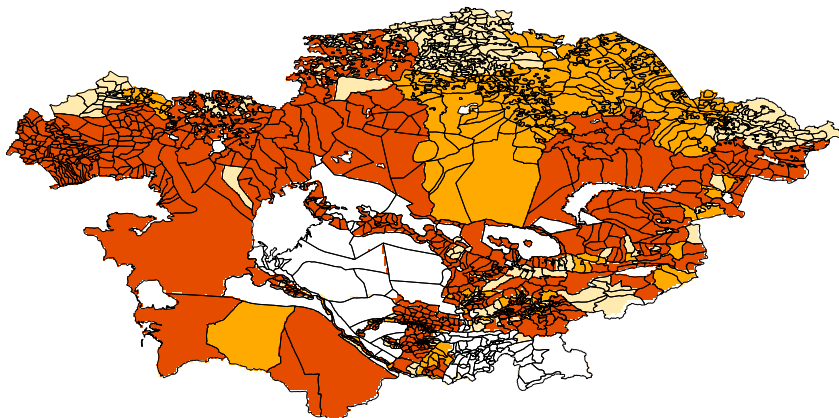
- Development of “socialist” nations (Official interpretation and Soviet historians) and “Affirmative Action Empire” (Carr, 1950; Slezkine, 1994; Martin 2001) ⇒
 - Both Political and Ethnic conflicts ↓

vs
- Colonial “Divide-and-Rule” policy in a different guise (Conquest 1986; Carrere d’Encausse 1988; Pipes 1997) ⇒
 - If the division in terms of ethnic groups: Political conflict ↓ while Ethnic conflicts ↑
 - However, if the aim was to divide a potentially-dangerous large Muslim state into smaller ethnic states, then: Political conflict ↓ and no prediction on ethnic conflict

vs
- Superficial administrative readjustment (Haugen, 2003) ⇒
 - No change or ↑ in both Political and Ethnic conflicts

► [Back to main slides](#)

Distribution of the local share of titular nationals



Color coding: yellow: below 33%; orange: 33%-66%; red: above 67%

White spots on the map: Tajikistan and Karakalpakia have different title ethnicities depending on the considered admin level, Republic vs. sub-Republic

► Back to main slides

► No heterogeneity by the share of TN

No heterogeneous effects by share of titular national group

Dep. Var. =	Ethnic Conflict Index		Political Conflict Index	
	(1)	(2)	(3)	(4)
Post × Reform	-0.018* (0.009)	-0.019** (0.009)	-0.018 (0.012)	-0.015 (0.012)
Post × Control × Share Titular National	-0.005 (0.012)	-0.012 (0.010)	-0.011 (0.013)	-0.007 (0.007)
Post × Reform × Share Titular National	0.002 (0.001)	-0.003 (0.003)	-0.003 (0.011)	0.012 (0.009)
Post × Control × Non-Complier	0.028** (0.011)	0.033** (0.015)	0.035*** (0.007)	0.032*** (0.011)
Post × Control × Share Titular National × Non-Complier		0.032 (0.038)		-0.021 (0.044)
Post × Reform × Share Titular National × Non-Complier		0.005 (0.003)		-0.022 (0.021)
Post × Reform × Non-Complier		-0.003 (0.002)		-0.012 (0.010)
Cell & Year FEs, 1922 Rep. Trends	✓	✓	✓	✓
Mean Dep. Var.	0.02	0.02	0.03	0.03
R ²	0.23	0.23	0.29	0.29
Observations	497,848	497,848	497,848	497,848

Why no heterogeneity? In homogenous areas, there could have been less conflict before (?)

► Back to main slides

► Map: share of TN

► Heterogeneity by shares of nomads and TN

Heterogenous effects by size of nomadic population

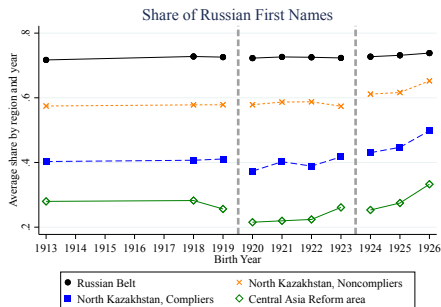
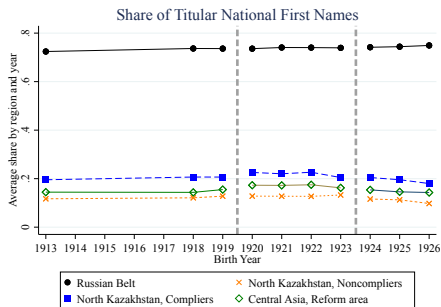
Dep. Var. =	Ethnic Conflict Index		Political Conflict Index	
	(1)	(2)	(3)	(4)
Post × Reform	-0.018* (0.009)	-0.018* (0.009)	-0.024** (0.011)	-0.025** (0.011)
Post × Control × Share Titular National	-0.005 (0.012)	-0.005 (0.012)	-0.012 (0.013)	-0.012 (0.013)
Post × Reform × Share Titular National	0.002 (0.003)	0.007*** (0.002)	0.000 (0.004)	0.006** (0.003)
Post × Reform × Less than 1/2 are Nomads	-0.002 (0.002)	-0.002 (0.002)	0.003 (0.003)	0.003 (0.002)
Post × Reform × Share Titular National × × Less than 1/2 are Nomads		-0.017** (0.006)		-0.017*** (0.005)
Post × Control × Non-Complier	0.028** (0.011)	0.028** (0.011)	0.035*** (0.007)	0.035*** (0.007)
1922 Rep. Linear Trends	✓	✓	✓	✓
Mean Dep. Var.	0.02	0.02	0.03	0.03
R ²	0.23	0.23	0.23	0.23
Observations	434,603	434,603	434,603	434,603

- Reform worked in places, where title nations dominate and population is already mostly settled population
- Backlash to the reform in places, where title nations dominate, but they are mostly nomadic

WWII Draftees: Classifying Ethnic Names

- ① Geolocate draftees birth place within 1939 Soviet districts
 - Match 1926 subdistrict centroids to 1939 districts; adjust by hand districts that straddle 1920s republics (only 5 of them)
- ② Obtain probabilities that a particular name belongs to a particular nation:
 - Feed first and last names separately into Forebears Onograph API
 - Bottom quartile of sum of top ten probabilities is 96% for first names and 99.5% for last names
 - For names that do not appear in Forebears, assign probabilities of similar names if similar enough
 - Normalize by sum of all Soviet titular nationalities
 - Use similar approach using Memorial data on ethnicity (authors' calculation)
 - Using Memorial data, correct Forebears' probabilities that reflect large stock of CA migrants in Russia today
- ③ Assign Titular National, Russian, and Other CA Titular National
 - Define binary measure based on threshold of 80%

Raw trends of the use of TN and Russian first names



► Back to main slides

A Bird's Eye View from the 1926 and 1939 Censuses

	Total Population		Urbanization Rate		Literacy Rate	
	1926	1939	1926	1939	1926	1939
Uzbekistan SSR	4,565,432	6,282,446	22%	23%	11%	68%
Turkmenistan SSR	998,154	1,253,985	14%	33%	13%	67%
Tajikistan SSR	1,032,216	1,484,440	10%	17%	4%	72%
Kyrgyzstan SSR	1,001,697	1,485,091	12%	18%	15%	70%
Kazakhstan SSR	6,073,979	6,151,102	9%	28%	23%	76%

- Annual population growth excluding Kazakhstan is 2.5%
- Kyrgyzstan's growth in urbanization is twice as fast as in Sub-Saharan Africa (Turkmenistan is nearly four times)
- Literacy rate in Sub-Saharan Africa below 67% today starting at about 10% in 1950s-60s.

[▶ Back to main slides](#)