



Colonization and development: The long-term effect of Russian settlement in the North Caucasus, 1890s–2000s



Timur Natkhov*

National Research University Higher School of Economics, Center for Institutional Studies, Russian Federation

ARTICLE INFO

Article history:

Received 20 December 2013

Revised 15 September 2014

Available online 22 September 2014

JEL classification:

N13

N33

O1

O15

Keywords:

Colonization

Social structure

Human capital

Long-run development

Natural experiment

ABSTRACT

Natkhov, Timur—Colonization and development: The long-term effect of Russian settlement in the North Caucasus, 1890s–2000s

This paper exploits differences in the proportion of Russian settlers in the North Caucasus in the nineteenth century to estimate the effect of colonization on long-term development. The identification strategy relies on the fact that the primary purpose of Russian colonization was to protect the country's access to warm-water ports. Therefore, settlement varied depending on the proximity to the Black Sea coast. Instrumenting the share of settlers by the distance to the coast, I show the positive impact of settlement on literacy among the indigenous population with long-term effects on income, educational attainment and the quality of local governance. To insure the validity of the instrument, I conduct a placebo test that shows that distance to the coast does not predict literacy and income in the South Caucasus, where Russians had no strategic interest in protecting the coastline. The mechanisms of influence include administrative integration, school building policies and social structure. *Journal of Comparative Economics* 43 (1) (2015) 76–97. National Research University Higher School of Economics, Center for Institutional Studies, Russian Federation.

© 2014 Association for Comparative Economic Studies. Published by Elsevier Inc. All rights reserved.

1. Introduction

What are the effects of colonial rule on long-run economic development? This question has generated significant debate in the empirical literature in recent years. Some scholars argue that the effect of colonization depends on the institutions established by the colonizer in the colony (Engerman and Sokoloff, 1997, 2002; Acemoglu et al., 2001, 2014). Others emphasize the importance of settlers' human capital, missionary activity and school building policies in shaping the subsequent development of a colony (Glaeser et al., 2004; Easterly and Levine, 2012; Woodberry, 2012; Wantchekon et al., 2013). Most of these studies exploit one important event – European colonization of the Americas, Asia and Africa – as a natural experiment of history that helps to identify the sources of long-run economic development (Nunn, 2009; Diamond and Robinson, 2010). In this regard, the case of the Russian Empire remains relatively unexplored. However, Russian history offers a perfect laboratory for testing different development theories using detailed data.¹

* Address: Shabolovka Street, 26, Office 5308, Moscow, Russian Federation.

E-mail address: timur.natkhov@gmail.com

¹ The expansion of the Russian state from a small Moscow principality into one of the largest empires in the world is comparable in speed and magnitude with the expansion of the British, Spanish and French Empires. However, unlike the European powers, Russia was a continental Empire (much more like the Habsburg and Ottoman Empires). Contiguous land borders with colonized people created different incentives for the metropolis and, as a result, a different set of institutions and policies in the colonies.

In this paper, I present a historical case – the colonization of the North Caucasus region by the Russian Empire in the late nineteenth century – that provides an opportunity for the empirical investigation of the effect of settlement on local populations and the subsequent regional development.

Based on a unique dataset of thirty North Caucasian districts collected from various historical records and the 1897 census, I show that Russian colonization altered the social structure of Caucasian communities and created new educational institutions, which resulted in human capital accumulation among the native population. In a preferred specification, a 10% increase in the Russian-speaking population is associated with a 5% increase in the literacy of the indigenous population. Moreover, this effect had long-term consequences for the educational attainment of the local population that persist today, in addition to affecting income levels, birth rates and the quality of local municipal governance in the districts.

Given the observed robust correlations between Russian settlement and long-term growth, the natural question is whether settlers selected districts that were already more literate and developed. More formally, settlement may be correlated with unobserved district characteristics that lead to biased estimates of the settlement effect. Using proxies for initial development, I show that this is not the case. Specifically, I show that there is no correlation between the initial population density in 1865 (before massive colonization started) and the share of Russian settlers thirty years later, in 1897. Similarly, I show that there is no correlation between the literacy of the local population in non-Russian languages and the share of Russian settlers. Thus, Russian settlers did not select for more dense and literate districts.

If initial development was not important for settlers, what was determining the pattern of settlement? Historical evidence shows that the primary purpose of colonization for the Russian Empire was to protect its southern frontier from its constant rival – the Ottoman Empire. To protect the border from the Ottomans in the North Caucasus, it was sufficient to control the Black sea coastline because the remaining border is protected by the mountains. Thus, the Russians' progress into the North Caucasus followed a clear geographical pattern – western districts closer to the Black Sea were much more extensively settled than the eastern ones close to the Caspian Sea. This difference resulted in a substantial cross-district variation in the share of Russian settlers that was exogenous to the local population. Indeed, the only variable that predicts the share of Russian settlers in the district is distance to the Black Sea coast – this correlation is very high (-0.82) and statistically significant at a less than 1% level. I pursue three strategies to show that distance to the coast is a valid instrument for Russian settlement.

First, I include distance to the coast directly into my baseline equation to see whether it has an independent effect on the literacy of the indigenous population. I find that when controlling for the share of Russians settlers, distance to the coast becomes insignificant.

Second, I explore whether distance to the coast has an impact on local development before Russian colonization. Proximity to the sea provides trade opportunities that encourage the accumulation of human capital and economic development.² Thus, the communities that live closer to the coast are more likely to be more developed. The results show that in the case of the North Caucasus before Russian colonization, distance to the coast does not predict the initial development of local communities as captured by population density and literacy in non-Russian languages.

Third, and most important, I conduct a placebo test. If distance to the coast affects the literacy of local people only through Russian settlement (i.e., if the exclusion restriction is satisfied), then there should be no relationship between distance to the coast and literacy in regions where Russians settled in inland areas and not on the coast. The unique geography and history of Caucasus colonization provides an opportunity for such a comparison. In the South Caucasus, separated from the North by the mountain range, proximity to the sea was of less military importance for Russians because this area has a land border with the Ottomans and also the Persians. Thus, the pattern of settlement in the South was very different from that in the North – Russians settled along the southern land border and not along the Black Sea coastline. Examining South Caucasian districts, I found no correlation between distance to the coast and the literacy of the indigenous population in 1897. However, there is a strong correlation between the share of Russians and the literacy of the indigenous population in the South as there is in the North. Even controlling for the population density, literacy of indigenous population in non-Russian languages and a number of geographical variables the magnitude of the effect is large and comparable with that in the North. Moreover, the share of Russian settlers in 1897 explains about 40% of the variation in income per capita today in a sample of South Caucasian districts, while distance to the Black sea coast is statistically insignificant.

The results of these tests indicate that the identifying assumption is plausible – in the North Caucasus distance to the coast affected the literacy of the indigenous population through Russian settlement and not through omitted variables. Similar to OLS coefficients, the IV coefficients are positive and significant, which suggests that an increased proportion of Russian settlers in a district positively affected literacy rates among the indigenous populations in the late nineteenth century.

The settlement effect had long-term consequences for local development. Districts that had higher shares of Russian settlers in the late nineteenth century today have higher income per capita, higher educational attainment among the indigenous populations, lower fertility rates and higher-quality municipal governance. Overall, the evidence supports the impact of Russian settlement on long-term development of regions in the North Caucasus.

Finally, I explore several potential mechanisms of influence – school building, social structure and infrastructure. In an attempt to integrate the Caucasus into the imperial administrative system, the Russian government established new

² A large body of literature beginning with [Gallup et al. \(1998\)](#) empirically documents the importance of geography and, particularly, proximity to the sea for international trade, transportation costs and long-run economic growth.

institutions and policies. The school building policy increased the number of state primary schools from 13 to 716 in fifty years. The policy was more intense in regions with a higher share of Russians; thus, it benefited the natives who lived closer to Russian settlements in the first place. There was also a growing trend to teach natives only in the Russian language. The data show that in the 1880s, only half of primary schools taught their students exclusively in Russian; the other half combined the local and Russian languages because there was a shortage of Russian teachers. By the end of the nineteenth century, however, more than 90% of schools taught their students only in the Russian language.

I also show that in settled districts, Russians constituted a major part (up to 80%) of core middle-class occupations that were essential for human capital transmission – education and public health. Thus, colonization was an event that changed not only the ethnic composition of the population but also the social and occupational structures in the region, resulting in divergent development paths for more settled and less settled districts. Also the number of civil administrators per native population was systematically higher in districts with a higher share of Russian settlers, indicating a greater capacity of the Russian state to implement new rules and regulations in the settled districts. At the same time railroads construction does not seem to have an effect.

This paper adds to a growing body of literature on the importance of history for economic development.³ Beginning with the seminal contributions of [Engerman and Sokoloff \(1997\)](#), [La Porta et al. \(1997\)](#), and [Acemoglu et al. \(2001\)](#), a significant body of empirical literature has argued that historical events and colonization patterns have had long-lasting effects on current development. The primary channels of this persistence are institutions and culture ([Nunn, 2008](#); [Nunn and Wantchekon, 2011](#); [Tabellini, 2008](#); [Guiso et al., 2008](#); [Dell, 2010](#); [Zhuravskaya et al., 2013](#)). A parallel line of research explores the effects of specific colonial policies on long-term economic development. [Huillery \(2009\)](#) analyses the impact of colonial rule across districts within French West Africa and finds a positive relationship between early colonial investment in education, health care and infrastructure and current levels of schooling, health outcomes and access to electricity and water. The persistence is explained by path-dependency of public investments – later investments continued to be located in areas that had many of them already. [Wantchekon et al. \(2013\)](#) using a unique micro-level data find a positive effect of school construction in colonial Benin. Comparing treated and untreated people in the same village they document large effects of education on living standards for the first generation of students, as well as for their descendants. They also find large village-level and extended family externalities. In the case of South America [Droller \(2013\)](#) finds that the areas of Argentina with historically higher shares of Europeans in the population currently have significantly higher per-capita GDP and educational levels. Similarly, [Rocha et al. \(2013\)](#) found that high-skilled European settlers in the state of San Paulo in Brazil fostered the long run development of settled regions through improvements in early levels of human capital. All these results are consistent with [Glaeser et al.'s \(2004\)](#) argument that European settlers brought their human capital with them to the colonies, thereby driving economic growth and institutional change.

This study also documents the positive long-term impact of settlers' human capital and school building policies adding to the literature the case of North Caucasus colonization. The study also empirically documents an additional channel of influence besides education – social structure. Since Russians were concerned with integration of the colonized territories into the imperial administrative system, these channels can be seen as complements rather than substitutes. To the best of my knowledge the only one paper that empirically investigates the impact of colonization in the context of Russian Empire is [Shubina et al. \(2014\)](#). They study the case of Central Asian settlement in the early twentieth century and find that distance to Russian settlements is a key determinant of agricultural technology adoption among Kazakh nomadic pastoralists. The effect is stronger for wealthier and less mobile Kazakh families with pasture land more suitable for sedentary agriculture.

The advantage of this paper is that the history of Caucasus colonization provides a unique opportunity to identify the settlers' effect by comparing outcomes in the North and South Caucasus – areas with very different settlement patterns. The focus on one relatively small region with a common history also contributes to the *ceteris paribus* condition, which is difficult to imitate in a cross-country or even a cross-regional setting.

The remainder of the paper is structured as follows. In the next section, I provide a short historical overview of the Caucasus region and the causes and manner of colonization by the Russian Empire. In Section 3, I describe the data used to estimate the effects of colonization. Section 4 documents the correlations in the data, and Section 5 addresses causality issues. Section 6 explores the mechanism of causal effect, and Section 7 examines the long-term consequences of Russian settlement. Section 8 concludes the paper.

2. Historical background

2.1. General information

The North Caucasus is the northern part of the Caucasus region between the Black and Caspian Seas, lying within European Russia. The region is characterized by enormous ethnic and language heterogeneity. This relatively small territory contains populations speaking 46 different languages in 8 language families ([Comrie, 2008](#)).

Historically, this region was populated by indigenous people organized mostly as tribal societies. Politically, the region has never been integrated and has never developed a modern centralized nation state, unlike those regions south of the Caucasian range in Georgia, Armenia and Azerbaijan. Located on the peripheries of the Ottoman, Persian, and Russian

³ See [Nunn \(2014\)](#) for a recent and very comprehensive review of this topic.

Empires, the region has been an arena for political, military, and cultural rivalries for centuries. The Ottomans were able to control parts of the northwest Caucasus until the late eighteenth century, largely through their ally, the Crimean Khan. At various times, the Persians were in control of Dagestan in the northeastern Caucasus. By the mid-nineteenth century, however, the Russians had succeeded in turning the Caucasus from a contested frontier zone into the borderland of the Russian Empire.

2.2. The purpose of colonization

To understand the Russian Empire's primary motivation for colonizing the Caucasus, it is important to place this event into a broader historical context. Historical studies show that geopolitical and military interests were the most important determinants of colonization and the nature of settlements. For instance, [Rhineland \(1975\)](#), in his study of Russia's imperial policy, argues that

"The annexation of the Caucasus was largely determined by [military] strategy, and strategists cared little for the peculiarities of Caucasian life. <...> Geography and chronology were perhaps the most important factors in any particular case [of the incorporation of each non-Russian area into the Empire]. The Caucasus faced Russia on the north, Turkey on the west, and Persia on the east. During the eighteenth century, as Russia extended her southwestern border with Turkey to the Crimea and the northern shores of the Black Sea, the Caucasus took on strategic importance as the eastern extremity of that border".

[[Rhineland, 1975](#), p. 218.]

Similarly, [Khodarkovsky \(1999\)](#) writes that

"The Russian colonial tool kit may not have been substantially different from the one used by the western colonial empires, but most of the time it was used for a different purpose. Unlike the European colonial projects in the Americas and Asia, which were predominantly driven by commercial interests, the Russian expansion in the Caucasus throughout the period was motivated primarily by the government's geopolitical concerns".

[[Khodarkovsky, 1999](#), p. 398.]

One of the central "geopolitical concerns" for the Russian Empire was the security of its southern borders and access to warm-water ports. The primary rival for these coveted regions for many years was the Ottoman Empire ([Allen and Muratov, 1953](#)). The Russian-Turkish wars extended into some of the longest conflicts in modern history. For two centuries, from 1700 to 1900, there were eight significant military conflicts between Russia and Turkey, one every 25 years on average. The wars were mostly sparked by Russia's attempts to establish warm-water ports on the Black Sea and to gain control of the Dardanelles and the Bosphorus straits to retain access to Mediterranean trade routes. Russia's victory in the 1828–1829 war resulted in the Treaty of Edirne, which gave Russia most of the eastern shore of the Black Sea ([Encyclopedia Britannica, 2013](#)). These new territories allowed Russia to move its southern frontier closer to the lands of local people – the Empire began its expansion into the North Caucasus using "heavy colonial machinery – the military, bureaucracy, missionaries, settlers, courts, and schools – to begin altering the traditional geographical and social landscape" ([Khodarkovsky, 2008](#)).

Russian colonization strategies ranged from violent military conquest to co-opting local elites, dispatching bureaucrats, merchants, and priests to the newly built towns, and encouraging peasants from neighboring Russian regions to settle and farm the land. The latter method has been the primary colonization path since the 1860s, when major military operations were completed. Russia's strategic goals were to integrate the region and its peoples into the Russian Empire ([Khodarkovsky, 1999](#)).

The historical evidence suggests that the arrival of new settlers in the North Caucasus region encouraged local populations to trade with Russian settlers. The natives provided livestock, sheepskins, clothes, leather goods, wood and lumber from upland forests. Russian settlers provided salt, metals and manufactured goods imported from Europe and other parts of Russia ([Barrett, 1995, 1999](#)). These mutual interests created a "frontier exchange economy"⁴ with intense cross-cultural interaction. Gradually, the indigenous people learned to speak Russian and the settlers adopted some local customs.⁵ Later, the Russian language formed the basis for local alphabets because most of the local languages were unwritten. Today, Russian serves as the region's lingua franca and is spoken by nearly everyone.

2.3. The North and South Caucasus: different colonization strategies

The South Caucasus (also referred to as Transcaucasia⁶) is a region located on the southern part of the Greater Caucasus Mountain range. Today, three independent states – Georgia, Armenia and Azerbaijan – are located there. All of them were part of the Russian Empire and later the Soviet Union.

⁴ This expression is coined by Daniel H. Usher, Jr. in "Indians, Settlers and Slaves in a Frontier Exchange Economy". Chapel Hill: University of North Carolina, 1992.

⁵ The most notable examples of cultural exchange were the adoption of local clothes and weapons by Russians. The traditional Caucasian dress (the *papakha* and *cherkeska*) – the tall fur hat and long tunic with rows of cartridge holders on the chest – eventually became the stereotypical image of the "Cossack". Further, high-quality swords (*shashka*) and daggers were supplied to Russian settlers by the natives ([King, 2008; Tolstoy, 1863](#)).

⁶ A common Russian term is *Zakavkaz'e*, which literally means "the area beyond the Caucasus Mountains".

The incorporation of this region into the Empire began with the annexation of the Georgian Kingdom in 1801. Over the next three decades, Russia sought to expand its territory south of the Caucasus at the expense of the Ottomans and the Persians. Although the purpose of colonization was the same as in the case of the North Caucasus – protection of the southern frontier – the methods were different. In the North Caucasus, Russian settlement was very extensive (by 1897, approximately 67% of the population were Russians) and was concentrated on the Black Sea coastline. In the South Caucasus, the settlement was limited (only 5% of the population in 1897) and concentrated along the border with Ottomans and Persians. Fig. 15 clearly shows the differences between the Russian colonial strategies for the North and South Caucasus. These differences were primarily determined by two factors – transportation costs and the geography of the border with rival empires (by sea or by land).

The North Caucasus is contiguous to European Russia. There are no natural barriers (high mountains or large rivers) between European Russia and the foothills of the North Caucasus. Thus, the transportation costs for Russian settlers were low. Conversely, the South Caucasus is separated from European Russia by the Caucasus mountain range. For this reason, Russians never settled in the South Caucasus in any great numbers, and the few who settled were mostly attached to the military or civil administration.

Proximity to the sea was of more strategic importance north of the Caucasus mountain range than south of it. To protect the border from the Ottomans in the North Caucasus, it was sufficient to control the Black sea coastline because the remaining border is protected by the mountains. However, the South Caucasus has a land border with the Ottomans and the Persians. There are no natural barriers and “there was nothing to stand in the way of cavalry galloping out of eastern parts of Ottoman Empire or northern Persia” (King, 2008, p. 27). Therefore, the control of ports in the South was not as important as the control of the land borders. These two factors – transportation costs from European Russia and the nature of the border with the Ottomans – were at the heart of the differences in Russia’s colonization strategies in the North and South Caucasus.

The most important conclusions of this section for our research are that the Russian colonization of the Caucasus was driven by geopolitical and military concerns and that the settlement pattern was very different between the North and South Caucasus. These facts provide an opportunity to identify the effect of Russian settlers on the indigenous population.

3. Data

I construct a dataset using a number of sources. The primary data source is the Russian Imperial census conducted in 1897, the only nationwide census conducted in the Russian Empire. This census contains information on literacy, ethnicity, religious affiliation, occupation and many other variables for all regions (oblast) and districts (okrug) of the Empire, including five regions in the North Caucasus⁷ and six regions in the South Caucasus.⁸ The North Caucasian regions are subdivided into 30 smaller districts that represent my main sample. The South Caucasian regions are subdivided into 45 districts. I use data from these districts for the placebo test in Section 5.3.

The data from the Russian Imperial Census of 1897 allows distinguishing people by their ethnicity defined in terms of native language. For each of the 75 districts, the census provides data on total population, the number of Russian-speaking people, the number of people speaking one of the local languages and the number of literate people among both Russians and the indigenous population. From these data, I calculate the proportion of Russians by district and the literacy rates among both Russians and the indigenous population. I match this historical data with modern data sources on income and the quality of municipal management from Rosstat (the Russian Official Statistical Agency), with educational attainment and birth rates from the latest available Population Census 2002 and with various geographical characteristics, such as temperature and elevation.

Although the number of observations (thirty districts) appears small for statistical analysis, this database has one advantage. Table 1 reports summary statistics and shows enormous variation in almost every dimension – economic development (literacy, population density), ethnic composition (Russian settlers) and geography (elevation, district area). For instance, literacy among the indigenous populations ranges from 1% to 32%, the proportion of Russian settlers ranges from less than 1% to almost 99%, the literacy among Russian settlers ranges from 11.5% to 78.6% and population density ranges from 2.6 to 55.8 people per square kilometer. As the following empirical analysis will show, this variation provides a good opportunity to estimate the effects of settlement on regional development.

4. Russian settlement and the literacy of the indigenous population: OLS estimates

I begin by examining the relationship between Russian settlement and the literacy of the indigenous population in 1897. My baseline estimating equation is

$$\ln \text{Indig Lit}_i = \beta_0 + \beta_1 \text{Rus Share}_i + \beta_2 \ln \text{Rus Lit}_i + \gamma \mathbf{X}'_i + \varepsilon_i \quad (1)$$

⁷ The North Caucasus regions are Chernomorskaya, Kubanskaya, Stavropolskaya, Terskaya and Dagestanskaya oblasts.

⁸ The South Caucasus regions are Kутаисская, Тифлисская, Бакинская, Эриванская, Elisabethpolskaya and Karsskaya oblasts.

Table 1
Summary statistics.

	Mean	Std. deviation	Min	Max	N
<i>Historical data (1897 census and 1860s statistics)</i>					
Literacy of indigenous population (Russian language), %	5.23	8.33	0.10	32.00	30
Literacy of indigenous population (non-Russian languages), %	6.93	6.13	0.37	26.32	30
Russian settlers, %	47.17	41.21	0.06	98.70	30
Russian settlers' literacy, %	29.56	17.46	11.52	78.64	30
Population density in 1897, per sq. km	22.39	11.09	3.44	55.76	30
Population density in 1865, per sq. km	15.12	9.81	1.50	44.02	21
<i>Modern data (2002 census and Rosstat data)</i>					
Russians, %	42.48	40.01	0.08	93.77	30
Russians with degree, %	34.81	8.85	19.26	53.13	30
Indigenous people with degree, %	22.49	10.35	8.35	43.52	30
Income per capita, thousands of rubles	11.60	5.11	4.60	23.16	30
Birth rates, per thousand of population	15.64	5.16	9.00	28.70	30
Federal subsidies in local budget, %	73.59	23.33	24.07	97.33	30
<i>Geography</i>					
District area, thousands of sq. km	6.96	4.54	1.08	16.69	30
Latitude	43.56	1.35	41.28	46.42	30
Elevation, meters	477.33	480.76	−6.00	1658.00	30
Average temperature in January, Celsius	2.32	2.54	−1.34	7.97	30
Distance to the Black Sea, km	406	253	0	789	30
<i>South Caucasus districts (used for placebo test)</i>					
Literacy of indigenous population (Russian language), %	3.2	3.9	0.3	15.7	45
Literacy of indigenous population (other languages), %	4.8	3.6	1.0	15.9	45
Russian settlers, %	4.0	5.5	0.1	24.9	45
Russian settlers' literacy, %	39.3	13.5	1.7	64.7	45
Distance to the Black Sea, km	295	196	0	704	45

Table 2
Russian settlers and literacy of indigenous population in 1897.

	Dependent variable: Log literacy of indigenous population in 1897 (Russian language)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Russian settlers	3.665*** (0.450)	4.750*** (0.782)	4.733*** (0.822)	4.843*** (0.843)	5.137*** (0.886)	5.011*** (0.970)	5.076*** (0.997)
Log settlers' literacy		0.986 (0.590)	0.982 (0.603)	1.067* (0.620)	0.863 (0.648)	0.917 (0.677)	1.018 (0.723)
Log literacy of indigenous population (non-Russian languages)			0.0191 (0.235)	0.0298 (0.237)	−0.0378 (0.245)	−0.000961 (0.271)	0.0273 (0.282)
Population density				−0.0125 (0.0170)	−0.0190 (0.0181)	−0.0179 (0.0187)	−0.0196 (0.0194)
District area					−0.0659 (0.0623)	−0.0645 (0.0636)	−0.0760 (0.0693)
Elevation						−0.203 (0.575)	−0.427 (0.760)
Temperature							−0.0496 (0.108)
Constant	−5.450*** (0.280)	−4.612*** (0.570)	−4.553*** (0.929)	−4.176*** (1.069)	−4.190*** (1.066)	−3.886** (1.387)	−3.353* (1.823)
Observations	30	30	30	30	30	30	30
R-squared	0.703	0.731	0.731	0.737	0.749	0.750	0.752

Note: Standard errors in parentheses.

* $p < 0.1$.** $p < 0.05$.*** $p < 0.01$.

where $\ln \text{Indig Lit}_i$ is the natural log of the literacy rate among the indigenous population in district i , Rus Share_i is the share of Russian settlers in the district, $\ln \text{Rus Lit}_i$ is the log of the literacy rate among Russian settlers and \mathbf{X}_i is a vector of control variables that captures differences in districts' economic development, geography and climate.

Table 2 reports the estimation results for Eq. (1). The first column reports simple regression results without any controls. In column (2), I account for the human capital of the settlers, i.e., the literacy rate of the settlers. In both specifications, the estimated relationship between Russian settlement and the indigenous people's literacy is positive and statistically

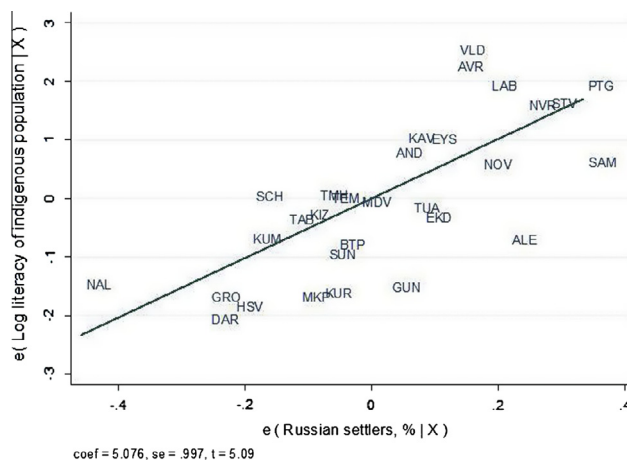


Fig. 1. Russian settlers and literacy of indigenous population in 1897.

Table 3

Determinants of Russian settlement.

	Dependent variable: Share of Russian settlers in 1897				
	(1)	(2)	(3)	(4)	(5)
Log literacy of indigenous population (non-Russian languages)	0.107 (0.0772)				0.0217 (0.0418)
Log initial population density		-0.223** (0.102)	-0.111 (0.0656)		
Elevation			0.1240 (0.1053)	-0.0766 (0.119)	-0.0927 (0.124)
Temperature			0.0456* (0.0231)	-0.0354* (0.0191)	-0.0332 (0.0198)
District area			0.0190 (0.0154)	0.0246** (0.0105)	0.0252** (0.0107)
Distance to the coast			-0.754*** (0.143)	-0.853*** (0.136)	-0.814*** (0.157)
Constant	0.798*** (0.247)	0.903*** (0.266)	0.911** (0.346)	0.991*** (0.178)	1.030*** (0.196)
Observations	30	21	21	30	30
R-squared	0.064	0.200	0.908	0.820	0.822

Note: Standard errors in parentheses.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

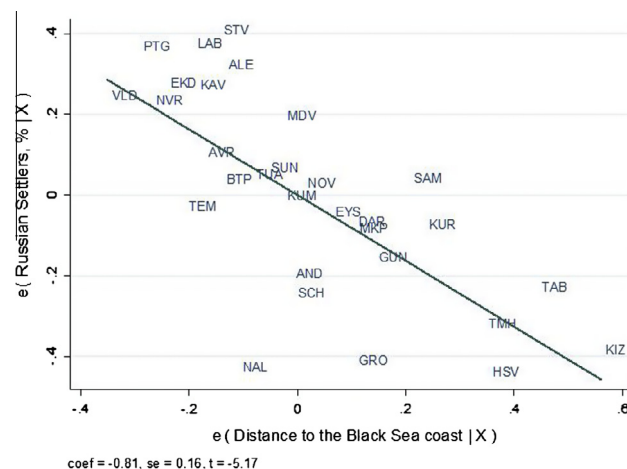


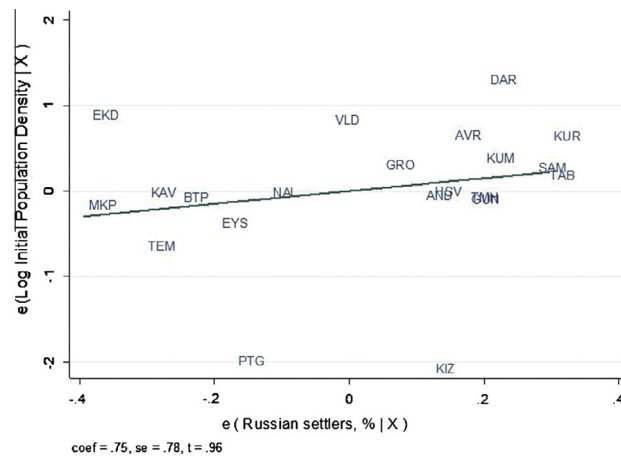
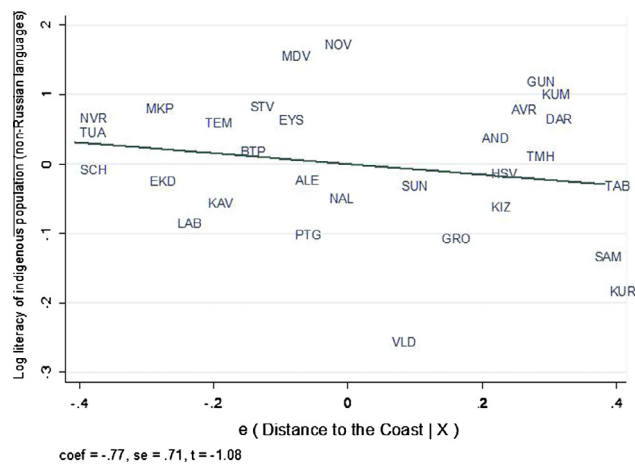
Fig. 2. Distance to the Black Sea coast and Russian settlement.

Table 4

Distance to the coast and initial development.

	Log literacy of indigenous population (Russian language)		Log initial population density in 1860s (3)	Log literacy of indigenous population (non-Russian languages) (4)
	(1)	(2)		
Distance to the coast	−5.121*** (0.934)	−1.447 (1.099)	0.750 (0.780)	−0.769 (0.715)
Russian settlers, %		2.996*** (0.675)		
Constant	−1.641*** (0.445)	−4.547*** (0.739)	2.107*** (0.423)	−2.747*** (0.340)
Observations	30	30	21	30
R-squared	0.518	0.721	0.046	0.040

Note: Standard errors in parentheses.

* $p < 0.1$.** $p < 0.05$.*** $p < 0.01$.**Fig. 3.** Distance to the coast and initial population density.**Fig. 4.** Distance to the coast and literacy of indigenous population in non-Russian languages.

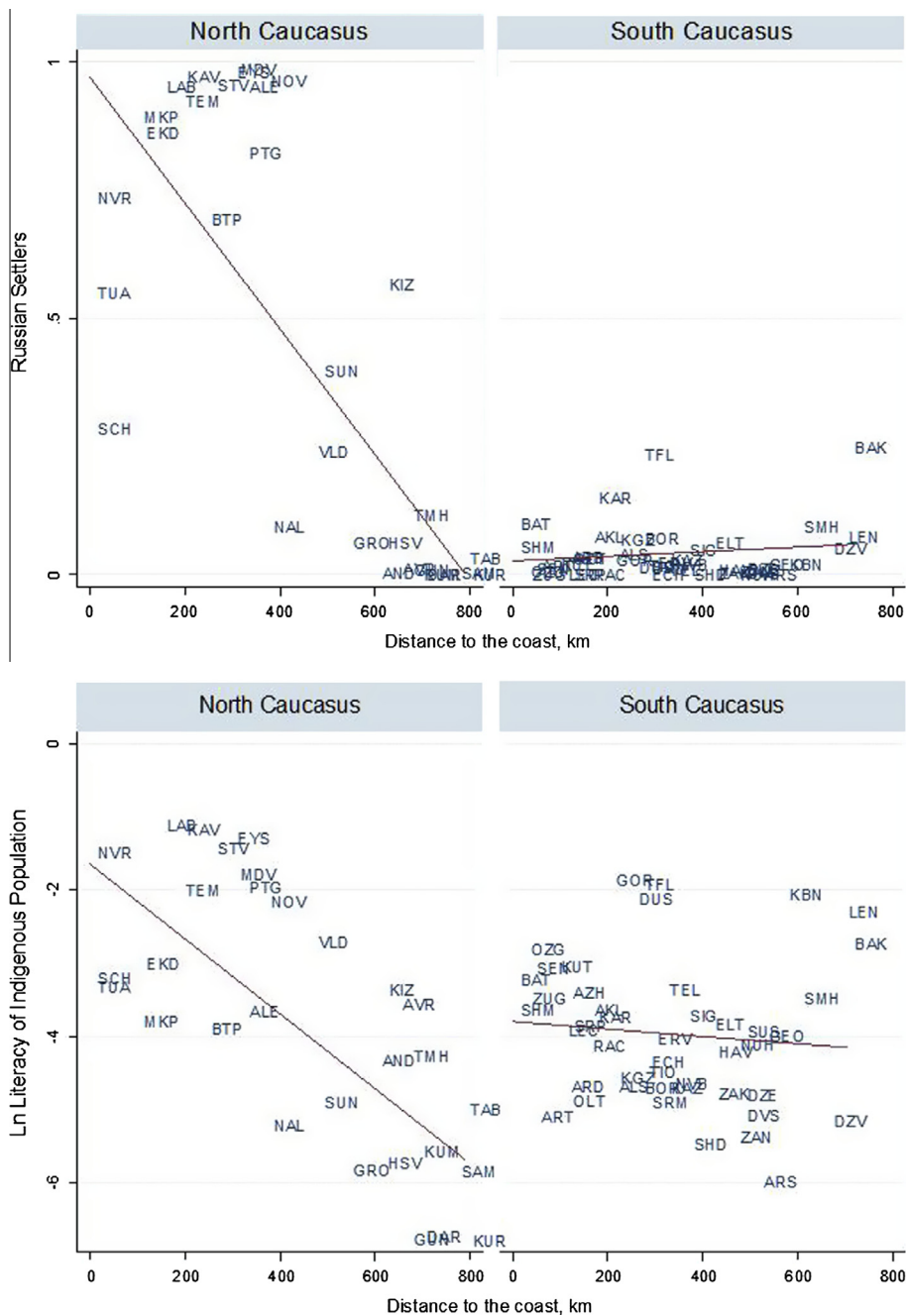


Fig. 5. Placebo test.

significant. In column (3), I control for the human capital of the indigenous people measured by literacy in non-Russian languages.⁹ In columns (4) to (7), I control for other possible covariates that account for economic development, geography and climate. Controlling for population density, district area, elevation of the largest town in the district and average temperature in January does not change the basic results.

The estimated magnitude of the relationship between Russian settlement and indigenous literacy is not only statistically significant but also very large. Calculating the standardized beta coefficients of the estimates, a one standard deviation increase in the share of Russian settlers is associated with an increase of 0.91–1.16 standard deviations in literacy rates

⁹ Most of the local languages were unwritten. Before Russian colonization Turkish and Arabic served as international languages in the North Caucasus. Therefore this variable captures mostly literacy in these two languages.

Table 5
Placebo test (South Caucasus districts).

	Russian settlers, %	Log literacy of indigenous population in 1897 (Russian language)			Log income per capita in 2005		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance to the coast	0.0446 (0.0420)	−0.500 (0.773)					0.222 (0.151)
Russian settlers, %			7.181*** (2.562)	5.362** (2.410)	1.827*** (0.455)	1.791*** (0.384)	1.575*** (0.401)
Log settlers' literacy				−1.158 (1.036)		−0.0297 (0.205)	−0.0213 (0.199)
Log literacy of indigenous population (non-Russian languages)				9.645** (4.175)		−3.091*** (0.720)	−2.492*** (0.810)
Population density				0.183* (0.0964)		0.158 (0.186)	0.163 (0.180)
District area				0.0668 (0.0954)		−0.0238 (0.178)	−0.101 (0.181)
Constant	0.03* (0.015)	−3.8*** (0.27)	−4.24*** (0.172)	−4.99*** (0.804)	8.12*** (0.036)	8.24*** (0.15)	8.188*** (0.148)
Observations	45	45	45	45	25	25	25
R-squared	0.026	0.010	0.155	0.40	0.412	0.753	0.758

Note: Standard errors in parentheses.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

among the indigenous population. Across thirty districts in the North Caucasus, a 10% increase in the Russian-speaking population was associated with an increase in literacy of approximately 5% among the indigenous population. Fig. 1 shows the estimation results from column (7).

5. Causality issues: IV approach

5.1. Determinants of settlement

To what extent was Russian settlement driven by self-selection? Did settlers choose districts with more literate indigenous populations and higher levels of economic development? To answer these questions, I exploit the earliest available data on population density in the districts.¹⁰ Before massive colonization started, in the late 1860s, the Russian government collected preliminary statistical information about the region.¹¹ These data are available for only 21 districts in my sample.

In addition, 1897 census provides information on literacy in languages other than Russian among indigenous populations. This variable is a good proxy for the human capital of indigenous people. Most of the local languages were unwritten, so two languages served as a lingua franca – Turkish and Arabic. These languages were used by merchants and religious authorities – a very small class of literate people in these traditional agrarian societies.

Both initial population density and literacy in non-Russian languages serve as proxies for the initial development level of districts. I use them as explanatory variables for Russian settlement. Table 3 reports the estimation results, in which I regress the proportion of Russian settlers in the district in 1897 on these development proxies.

Column (1) shows that the coefficient of the literacy rate in non-Russian languages is positive but statistically insignificant – there is no relationship between Russian's decision to settle and the literacy of the indigenous population. In column (2) I use initial population density as an explanatory variable for settlement. At this time, the coefficient is negative and statistically significant at a 5% level. It might be surmised that the Russians selected themselves into regions with lower population densities and, therefore, a lower level of development, but this effect disappears when I control for several geographical variables in column (3), which shows that there is no relationship between population density in the late 1860s and the proportion of Russian settlers thirty years later. Based on the correlations in Table 3 (columns 1–3), I conclude that Russian settlers did not choose settlement areas based on initial development level. This statistical finding is consistent with established historians' view of the nature of North Caucasus colonization (see Section 2.2).

¹⁰ It is well acknowledged in the literature that population density is a good proxy for economic development in the pre-industrial era because only societies with significant agriculture productivity were able to support dense populations. See Acemoglu et al. (2002).

¹¹ These data contain lists of villages and towns in each district, populations, numbers of households and other general information. See "Statistical Table of Caucasus Region in 1865" ("Statisticheskaya tablitsa Kavkazskogo kraya v 1865". Tiflis, 1866).

Table 6

Russian settlers and literacy of indigenous population in 1897 (IV estimation).

	Second stage Dependent variable: <i>Log literacy of indigenous population (Russian language)</i>			
	(1)	(2)	(3)	(4)
Russian settlers, %	4.175*** (0.611)	7.391*** (2.102)	7.998*** (2.384)	4.556*** (1.200)
Log settlers' literacy		2.640* (1.373)	2.134* (1.164)	0.122 (0.267)
Log literacy of indigenous population (non-Russian languages)			−0.262 (0.391)	0.204 (0.182)
Population density			−0.0381 (0.0264)	0.800 (0.970)
District area			−0.145 (0.0954)	−0.0566 (0.0651)
North Caucasus dummy				−1.554** (0.510)
Distance to the coast				0.146 (0.816)
Geography controls	No	No	Yes	Yes
Second stage <i>R</i> -squared	0.689	0.618	0.656	0.559
	First stage Dependent variable: <i>Russian settlers, %</i>			
	(1)	(2)	(3)	(4)
Distance to the coast	−1.226*** (0.202)	−0.558*** (0.216)	−0.791*** (0.297)	0.025 (0.119)
Log settlers' literacy		−0.447*** (0.100)	−0.121 (0.150)	−0.046 (0.035)
North Caucasus dummy				0.712*** (0.073)
Distance to the coast * North Caucasus dummy				−0.948*** (0.162)
Full set of controls	No	No	Yes	Yes
Observations	30	30	30	75
<i>R</i> -squared	0.567	0.750	0.836	0.852
AR-test <i>p</i> -value	0.0001	0.0001	0.0001	0.0362
FAR-test <i>p</i> -value	0.0030	0.0035	0.0015	0.0953
First stage <i>F</i> -statistics	36.69	40.61	16.09	55.19

Note: All regressions contain a constant term. Standard errors in parentheses.

** $p < 0.05$.

* $p < 0.1$.

*** $p < 0.01$.

Column (3) in Table 3 shows that the distance to the Black Sea coast is the strongest predictor of the share of Russian settlers across 21 North Caucasian regions. In column (4), I exclude the initial population density to increase the number of observations and show that geography alone explains 82% of the variation in Russian settlement in the full sample of thirty districts. Column (5) shows that when controlling for literacy and a number of geographical variables such as district area, elevation and temperature, the only variable that significantly (statistically and economically) predicts Russian settlement is the distance to the Black Sea coast. Fig. 2 shows the conditional correlation as estimated in column (5).

5.2. Is distance to the coast a valid instrument?

The validity of the instrument relies on two properties – it must be correlated with Russian settlement, and it must not be correlated with other characteristics that might affect the literacy of the indigenous population. While the first property is confirmed in Table 3, the second could not be explicitly tested. For instance, if proximity to the Black Sea was correlated with trade potential and potential returns to the accumulation of human capital, then distance to the coast could affect literacy directly and not via Russian colonization. In this case, distance to the coast cannot be used as an instrument. I conduct three statistical tests to show that it is very unlikely that the exclusion restriction is violated.

First, I directly include distance to the coast as a regressor in my baseline Eq. (1). If distance to the coast has an independent effect on literacy (beyond Russian settlement), then the coefficient must be statistically significant. However, distance to the coast does not predict the literacy of the indigenous population once I control for the share of Russian settlers (column 2 in Table 4).

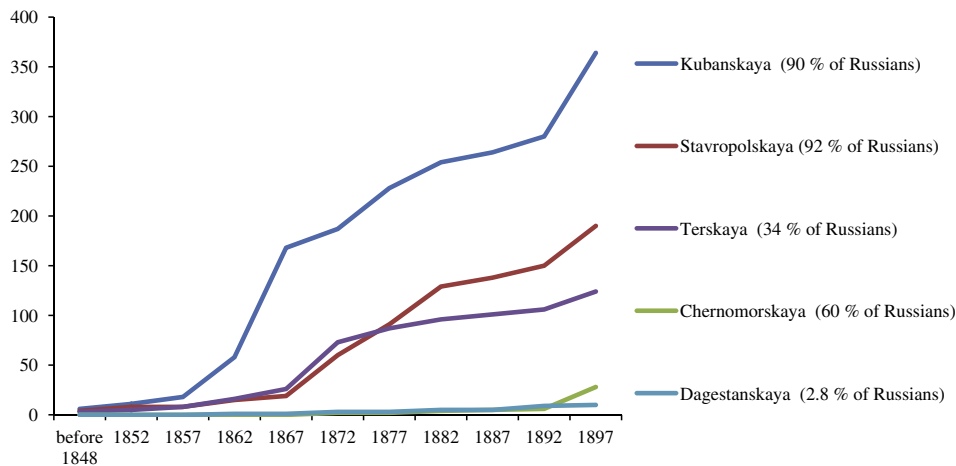


Fig. 6. Schools construction in the North Caucasus, 1848–1897 (absolute number of state primary schools).

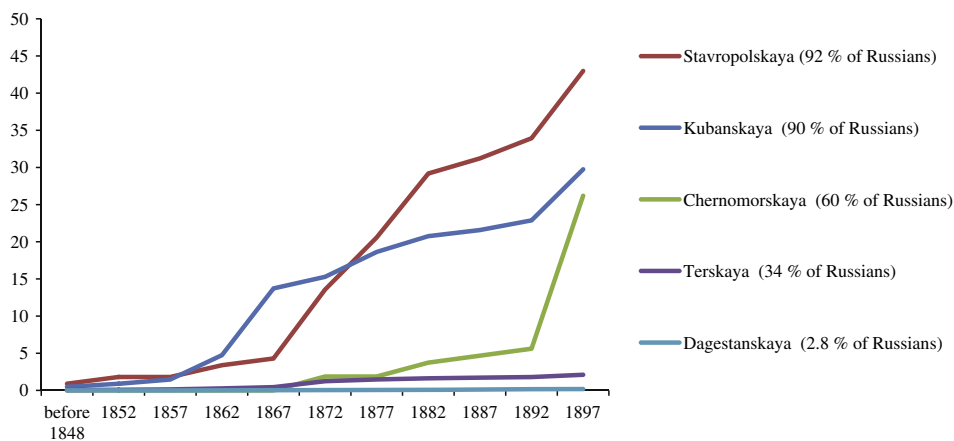


Fig. 7. Schools construction in the North Caucasus, 1848–1897 (number of state primary schools per 10,000 of indigenous population).

Second, I test whether distance to the coast is correlated with initial development using data on initial population density and literacy in non-Russian languages. As columns (3) and (4) in Table 4 show, there is no correlation between the initial development level of North Caucasian districts and distance to the coast (see also Figs. 3 and 4). Distance to the coast became important for local development only after the Russians began to settle the area.

Third, and most importantly, I conduct a placebo test by estimating the effect of distance to the Black sea coast on literacy in the South Caucasus. I exploit the fact that the geography of Russian settlement was very different in the North and South Caucasus. In the South Caucasus, the proximity to the sea was of less strategic and military importance; Russians settled along the southern land border with the Ottomans and Persians and not along the Black Sea coastline. Thus in the South Caucasus, distance to the coast could not have a noticeable effect on the literacy of indigenous population if the only channel of influence is Russian settlement. At the same time Russian settlement, regardless of its geographical location, should predict the literacy of indigenous population and other development outcomes.

Indeed, Fig. 5 shows that in the South, there is no correlation between the distance to the Black Sea on the one hand and Russian settlement and the literacy of the indigenous population on the other hand (see also columns 1 and 2 in Table 5). However, there is a strong correlation of the share of Russians with the literacy of the indigenous population in the South, just as there is in the North (column 3 in Table 5). Even controlling for population density, literacy in non-Russian languages and geography, the magnitude of the effect is large and is comparable with that in the North (column 4 in Table 5). Moreover, column 5 of the same table shows that the share of Russian settlers in 1897 explains about 40% of the variation in income per capita today in the sample 25 South Caucasus regions.¹² In column 6 I add the same set of controls; the results do not

¹² The data on income per capita in the regions of South Caucasus are from Gennaioli et al. (2013). I could successfully match only 25 regions from my 45 regions sample because there were vast changes in administrative borders of new independent states of Georgia, Armenia and Azerbaijan.

Table 7

Mechanisms of the effect: school building.

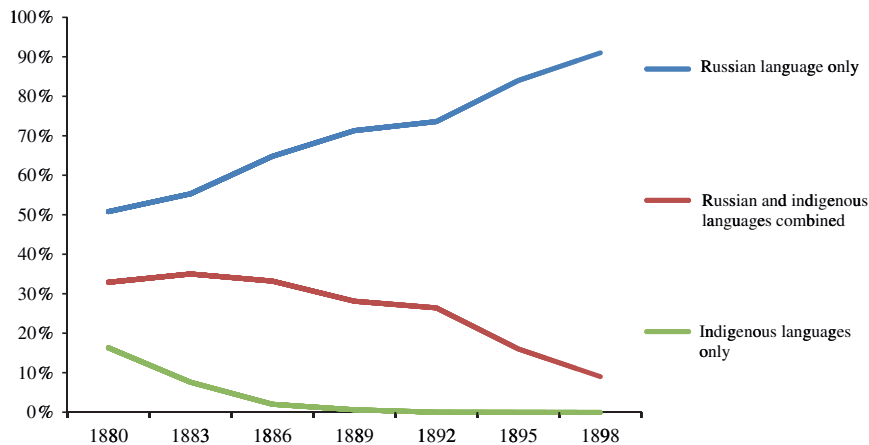
	Number of schools in 1897 (3)	Number of schools per 10,000 of population in 1897 (4)	Number of schools per 10,000 of indigenous population in 1897 (5)
Russian settlers	46.77*** (6.134)	2.353*** (0.810)	177.7* (97.36)
Log settlers' literacy	−2.890 (4.445)	−0.867 (0.587)	22.54 (70.55)
Log literacy of indigenous population (non-Russian languages)	−3.111* (1.735)	−0.0448 (0.229)	58.92** (27.54)
Population density	0.257** (0.119)	−0.0621*** (0.0158)	0.994 (1.895)
District area	1.615*** (0.426)	−0.197*** (0.0563)	−8.529 (6.768)
Elevation	0.0127** (0.00467)	−9.83e−06 (0.000617)	−0.135* (0.0741)
Temperature	0.559 (0.662)	0.225** (0.0874)	−23.04** (10.50)
Observations	30	30	30
R-squared	0.941	0.778	0.586

Note: All regressions contain a constant term. Standard errors in parentheses.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

**Fig. 8.** Schools by language of study in the North Caucasus, 1880–1898.

change – a one standard deviation increase in the share of settlers is associated with 0.62 standard deviations increase in income per capita in 2005. In column 7 I check whether distance to the coast has an effect on contemporary income; it is actually insignificant.

The results of this placebo test suggest that in the South Caucasus Russian settlements predict the literacy of indigenous population in 1897 and income per capita today. On the contrary, distance to the Black sea coast is unable to do so, because Russians did not settle on the coastline in the South. Thus the identifying assumption that in the North Caucasus distance to the coast affected the literacy of the indigenous population through Russian settlement and not through omitted variables is plausible.

5.3. IV estimation

According to historical and statistical evidence, I can interpret Russian settlement “as if a random” treatment that generates a natural experiment of history (Diamond and Robinson, 2010; Dunning, 2012). I use distance to the Black Sea as an instrumental variable for the share of Russian settlers and then use the predicted share of settlers as an explanatory variable for literacy in the main Eq. (1).

The results of the IV estimation are reported in Table 6. The first column reports estimates without control variables, the second includes the literacy of settlers, and the third includes development and geography controls. The first stage estimates

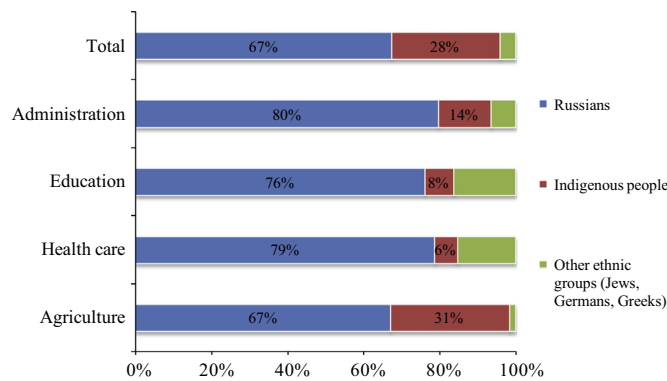


Fig. 9. Occupational structure by ethnicity in the North Caucasus, 1897.

Table 8

Mechanisms of the effect: social structure in 1897 and railroads.

	Employment in civil administration (1)	Employment in education (2)	Employment in health care (3)	Employment in agriculture (4)	Size of the middle class (5)	Year of first railway station (6)
Russian settlers, %	8.242** (3.214)	3.625** (1.626)	1.917* (0.961)	−21.36 (22.74)	19.08*** (6.013)	−28.91 (41.65)
Log settlers' literacy	3.931 (2.329)	0.644 (1.178)	0.430 (0.697)	−68.74 (95.46)	6.833 (4.357)	−11.70 (30.18)
Log literacy of indigenous population (non-Russian languages)	−1.066 (0.909)	−0.567 (0.460)	−0.202 (0.272)	2.513 (37.94)	−2.420 (1.701)	9.580 (11.78)
Population density	−0.0580 (0.0625)	−0.0165 (0.0316)	−0.0136 (0.0187)	−0.929 (2.616)	−0.233* (0.117)	0.0756 (0.811)
District area	−0.502** (0.223)	−0.180 (0.113)	−0.135* (0.0668)	6.399 (9.488)	−1.443*** (0.418)	−1.763 (2.895)
Elevation	−0.148 (0.245)	0.0262 (0.124)	−0.0375 (0.0732)	0.0112 (0.115)	−0.564 (0.458)	0.354 (0.317)
Temperature	0.222 (0.347)	0.189 (0.175)	0.0922 (0.104)	−26.46* (15.30)	1.146* (0.649)	−3.860 (4.493)
Observations	30	30	30	30	30	30
R-squared	0.447	0.351	0.396	0.348	0.686	0.428

Note: All regressions contain a constant term. Standard errors in parentheses.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

are reported in the bottom panel of the table. The coefficient for the instrument is negative, which suggests that the further the district is from the Black Sea coast, the fewer Russians settled there. The first stage F-statistics are also high in all specifications and cross the rule-of-thumb threshold of 10.

In column 4 I pool the North and South samples to increase the number of observations from 30 to 75. In this specification I instrument the share of Russian settlers with the interaction of the distance to the coast and North Caucasus dummy. The interaction term predicts settlement with p -value much less than 1%. The North Caucasus dummy separately is also significant and positive indicating a much more extensive Russian settlement in the North. On the other hand, distance to the coast separately becomes insignificant confirming that proximity to the sea was important only in the North.

It is worth noting that the coefficient of interest in the second stage is almost the same as in the OLS estimations for the subsamples of North and South districts (5.1 for North, 5.3 for South and 4.6 for pooled sample with IV). A test on the exclusion restriction developed by Berkowitz et al. (2012) and Riquelme et al. (2013) indicates the validity of the instrument. Overall, the IV results confirm the OLS estimates, with the IV coefficients slightly different than the OLS ones.

6. Mechanisms of the effect

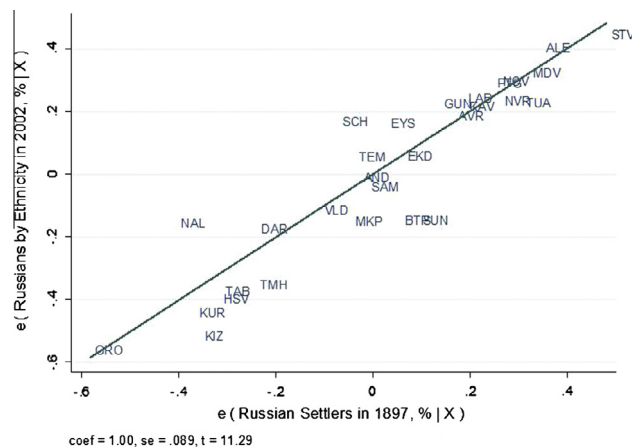
The central problem that the Russian rulers faced in the North Caucasus was the absence of centralized political authority. The region was a fragmented world of traditional societies with rudimentary political organizations. None of the areas possessed political elites who could exercise their monopoly of power upon a large territory and a large number of people. Thus,

Table 9

Long-term effect of Russian settlers. OLS estimates.

	Russian population in 2002 (1)	Log income per capita in 2010 (2)	Education among indigenous population (3)	Log birth rate in 2011 (4)	Share of federal subsidies in local budget in 2007 (5)
Russian settlers in 1897, %	0.955*** (0.106)	0.874*** (0.206)	0.126* (0.0726)	−0.945*** (0.143)	−0.504*** (0.122)
Russians with degree in 2002, %	0.467 (0.366)	0.837 (0.713)	0.306 (0.251)	−0.189 (0.495)	−0.781** (0.367)
Log literacy of indigenous population (non-Russian languages)	0.0242 (0.0344)	0.0790 (0.0671)	−0.0172 (0.0236)	0.0483 (0.0466)	−0.0216 (0.0379)
Population density	−0.00166 (0.00231)	−0.00704 (0.00450)	−0.000881 (0.00159)	0.00805** (0.00313)	0.00469* (0.00236)
District area	−0.0199** (0.00836)	−0.0469*** (0.0163)	0.00114 (0.00575)	0.0160 (0.0113)	0.0240** (0.00960)
Elevation	−0.000138 (9.79e−05)	−4.25e−05 (0.000191)	5.89e−05 (6.73e−05)	−0.000371** (0.000133)	2.64e−05 (0.000113)
Temperature	0.00471 (0.0127)	0.0398 (0.0248)	0.0253*** (0.00874)	−0.0244 (0.0172)	−0.0291* (0.0143)
Observations	30	30	30	30	30
R-squared	0.929	0.744	0.498	0.766	0.821

Note: All regressions contain a constant term. Standard errors in parentheses.

* $p < 0.1$.** $p < 0.05$.*** $p < 0.01$.**Fig. 10.** Russian settlement in 1897 and proportion of Russians in 2002.

for Russia the colonization of the Caucasus “was not so much a question of how to divide and rule as how to unite and absorb” (King, 2008, p. 38). The creation of the Caucasian Vicegerency (*Kavkazskoe namestnichestvo*) in 1844 was a first attempt at such unification.

The new institution was created by Emperor Nicolas I “to establish a solid structure of government in the Caucasus as soon as possible”, to “give more unity and speed” to the implementation of all matters regarding the region and finally to integrate the Caucasus into the imperial administrative system (Rhinelander, 1981). In the following decades, new administrative boundaries were created, new regulations issued on local government and completely new educational institutions were established that would make the inhabitants “speak, think, and feel in Russian”. In this section, I explore various channels through which this institution affected regional development.¹³

¹³ One of the potential channels could reveal itself before Russians began to build government institutions. Several tribes resisted colonization and Russians initiated military operations to pacify the natives. Although the military stage of colonization was mostly finished in 1850s (the paper deals with the period from the early 1870s, i.e. twenty years later), it could be the case that the effect on literacy is driven by the selection effect such that only the smartest out of indigenous population may have survived. I cannot completely rule out this possibility since there are no data on a number of killed native people. However, the data on mortality of Russian soldiers do not reveal any consistent pattern (Gizetti, 1901). Soldiers were equally likely to die regardless of the distance to the coast; other geographical and economic variables (for instance, initial population density) do not seem to have an effect as well.

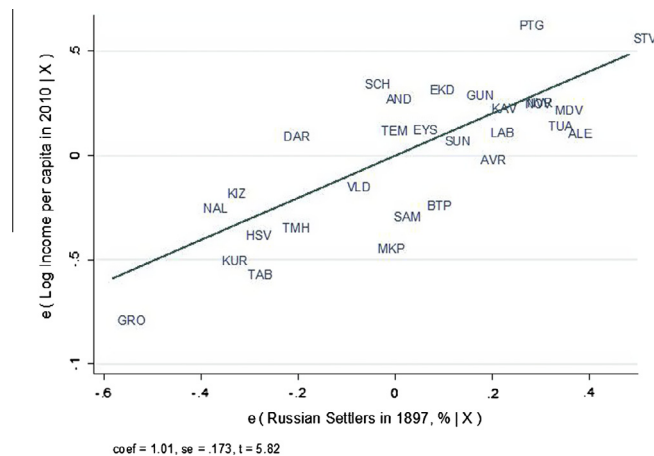


Fig. 11. Russian settlement in 1897 and income per capita in 2010.

6.1. Construction of primary schools

In 1848, four years after the establishment of the Vicegerency, an imperial law created the Caucasus school district, which brought the schools in the Caucasus regions into the regular Russian educational system. The primary advocate for the law was Count Mikhail Semenovich Vorontsov, who was a Vicegerent (*Namestnik*) of the Caucasus territory in 1844–1854. His intention was that education should play a preeminent role in the assimilation of native people into the Empire.¹⁴ For many progressive Russian imperial administrators, such as Vorontsov, the best assimilation policy was “to impress the natives with the benefits of Russian civilization by example, rather than by the force of blind assumption that the benefits were self-evident” (Rhineland, 1975, p. 232).

The school building policy was fairly successful. According to official statistics, there were only thirteen state primary schools in the North Caucasus in 1848 – the year the Caucasus school district was established. Fifty years later, by 1898, there were 716 state primary schools. There was a large geographical variation however – most schools were built in regions with extensive Russian settlement. As a result, the policy benefited the natives who lived closer to Russian settlements in the first place. Fig. 6 exposes this remarkable difference.

On the one extreme is Kubanskaya oblast (region), with 90% of Russians in 1897, where the absolute number of primary schools increased from 6 to 364. On the other extreme is Dagestanskaya oblast, with only 2.8% of Russians; here, there were no state primary schools in 1848 and only 10 schools in 1897.¹⁵

The data on the absolute number of schools can be misleading because more populous regions tend to have more schools. In Fig. 7, I normalize the absolute number of primary schools by the number of indigenous people in the region in 1897 to account for differences in population. This adjustment actually strengthens the result. For instance, Chernomorskaya region, with 60% of Russians, had only 28 state primary schools in 1897 and appears as an outlier in the previous figure. Adjusting for the number of indigenous population makes a significant difference – Chernomorskaya region becomes the third after Stavropolskaya and Kubanskaya, with 92% and 90% of Russians, respectively.

Although long time series data on school building is available only on a regional level, a cross section for 1897 is available on district level. In Table 7, I use these data to explore the determinants of school construction. I run the baseline regression (1) with the results of the school building policy as a dependent variable. This analysis shows that the share of Russian settlers in a district is the best predictor of the absolute number of schools in 1897 (column 3), the number of schools per 10 thousand of total population (column 4) and the number of schools per 10 thousand of indigenous population (column 5).

The other important feature of schooling as a channel of influence is the language used in teaching. At the early stages of the policy, only 50% of schools taught natives the Russian language. The other half of the schools combined the local and Russian languages or taught exclusively in the local language as there were not enough Russian teachers (Shonger, 1899). By the end of the nineteenth century, however, more than 90% of schools taught their students only in the Russian language (see Fig. 8).

¹⁴ Vorontsov was an exceptional administrator. Born in 1782 to the family of the Russian ambassador in Britain, he was reared in an English environment, attended Cambridge and moved to Russia only at age sixteen. During the Napoleonic Wars, he was a commander of the Russian occupational forces in France. At the age of forty-one, he was handed the governor-general's post in New Russia (today southern Ukraine). Under his administration, the city of Odessa was transformed from a small village to a thriving port and one of the largest cities of the Empire. In recognition of his achievements in 1844, Tsar Nicholas I granted him the title Vicegerent (commander in chief) of Caucasus with full and unlimited powers. It appears likely that Nicholas decided to create the Vicegerency only to persuade Vorontsov to take the burden of the Caucasus administration (Rhineland, 1990).

¹⁵ The long time series data on the number of schools are available only on a regional level. The data are from the official “Report of Caucasian school district on educational institutions in 1897” (“*Otchet popechitel'ya Kavkazskogo uchebnogo okruga o sostoyanii uchebnykh zavedeniy za 1897 god*”). Tiflis, 1898).

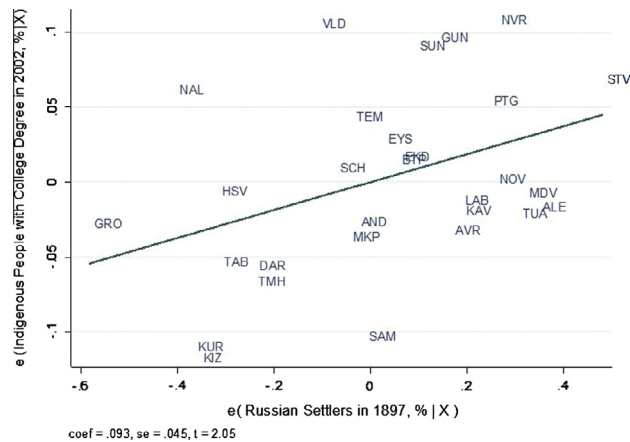


Fig. 12. Russian settlement in 1897 and indigenous population with college degree in 2002.

To sum up, the data on school construction show that regions with a higher share of Russian settlers experienced the most notable growth in the number of primary schools during the second half of the nineteenth century. There was also a growing trend to teach natives only in Russian and to neglect local languages as a school subject. Both factors contributed to the rise of literacy among the indigenous population, especially in regions with a higher share of Russians.

6.2. Social structure

As Russians were on average more literate than the indigenous people, they constituted a major portion of the workers in occupations that required high skills. Fig. 9 shows occupational structure by ethnicity in North Caucasian districts according to the 1897 census. Although Russians constituted 67% of the overall population, they were overrepresented in the education (76%) and health care (79%) sectors. In contrast, indigenous people constituted 28% of the overall population, but they only represented 8% of the education and 6% of the health care sectors. Instead, they were overrepresented in agriculture (31%).

In Table 8, I regress the shares of different occupations on the share of Russian settlers across the thirty districts. The results show that the share of Russian settlers predicts employment in administration, education and health care, but does not predict employment in agriculture. Controlling for geography and population density does not change the results. The magnitude of the coefficients is also very large – an increase of one standard deviation in the share of Russian settlers increases employment in education by 0.82 and employment in health care by 0.74 standard deviations.

Using the data on occupational structure I compute an integral variable – the size of the middle class, which includes employment in civil administration, education, health care and also trade. Column 5 in Table 8 shows that the share of Russian settlers is the best predictor of the size of the middle class across thirty North Caucasian districts in 1897.

In sum, by the end of the nineteenth century, Russians constituted not only a large share of the North Caucasus population, but they also constituted a major share of a key strata of society that was essential for the accumulation of human capital and long-run economic development.

6.3. Railroads

The other potential channel for the effect of colonization is infrastructure. Railroad construction was one of the drivers of industrialization and rapid economic development in late Imperial Russia (Gregory, 1994).

In column 6 of Table 8, I regress the year that the first railroad station was built in a district on a number of potential determinants. The results show that Russian settlement does not predict railroad construction. One of the possible explanations is that railroads in the North Caucasus were mostly built to connect European Russian with South Caucasus. For instance, the first line was built in 1875 from Rostov-on-Don to Vladikavkaz – the town in the center of the mountain range where the Georgian Military Highway begins. The goal was to lower transportation costs from European Russia to the administrative center of the Caucasus Vicegerency in Tiflis, Georgia.

Prior to the beginning of the twentieth century, two additional lines were built – to the Caspian Sea coast in the east and to the Black Sea coast in the west. However, only nine out of 30 districts were connected to the imperial railroad network by that time.

7. The long-run effect of Russian settlement

In this section I examine the long-run consequences of Russian settlement in the North Caucasus by linking 1897 data on settlement to modern data on income, educational attainment and other indicators of development. First, I document a

strong persistence in settlement – districts that happened to have higher proportions of settlers in the late nineteenth century have higher proportions of Russian population today. A simple correlation between the share of Russians in 1897 and in 2002 is 0.93. In Table 9, column (1), I take into account the current educational level of the Russian population and the same set of controls. The results are robust, which suggests that even one hundred years after the active phase of colonization, the population composition in the North Caucasian districts is almost the same (Fig. 10).

This persistence of Russian settlement resulted in divergent development paths for Caucasian societies. In column (2), I show that districts with higher shares of Russian settlers in 1897 have higher income per capita today even when controlling for the current educational level of the Russian population (Fig. 11). This is the effect on the entire population of the district because, unfortunately, the data do not allow distinguishing between the income of Russians and the income of indigenous population.

However, the 2002 Russian Federation Census provides this opportunity to account for educational attainment. From these census data, I calculate the share of the population with college and university degrees among Russians and the indigenous population based on self-reported ethnicity. In column (3), I regress the share of indigenous population with college or university degrees on the proportion of Russian settlers in 1897, controlling for the current educational level of Russian population and other district characteristics; I found a strong positive relationship (Fig. 12). Russian settlement affects the human capital of indigenous population in times of colonization and continues through the present day.

The human capital theory predicts that rising educational levels will be associated with decreasing fertility rates. Statistical evidence suggests that this is the case in almost every developed and developing country. Consistent with the theory, the evidence shows that in districts with higher proportions of Russian settlers (and with higher literacy rates) in 1897, the fertility rates today are significantly lower, even after controlling for the current educational level of the Russian population (column 4 and Fig. 13).

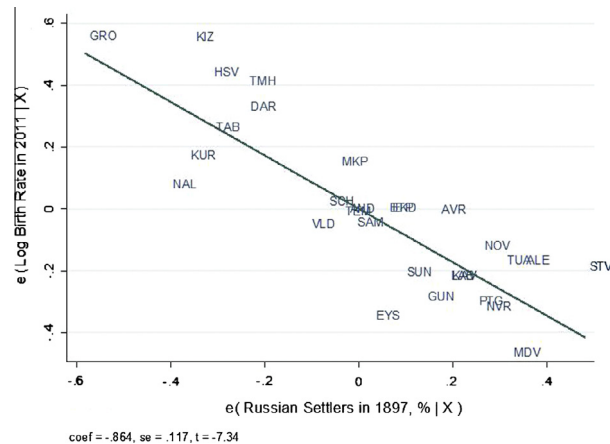


Fig. 13. Russian settlement in 1897 and birth rates in 2011.

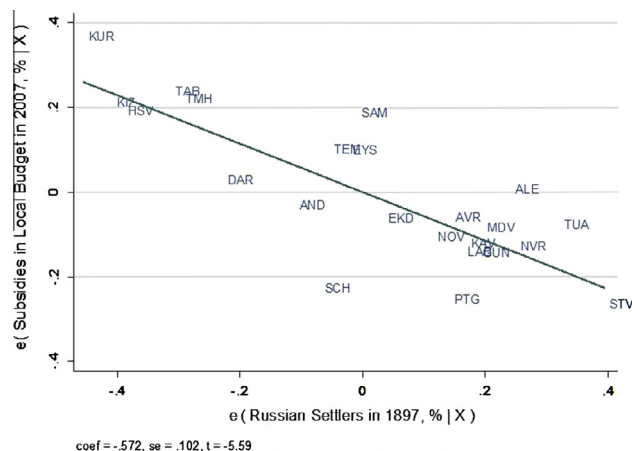


Fig. 14. Russian settlement in 1897 and share of federal subsidies in local budget in 2007.

Table 10

Long-term effect of Russian settlers. IV estimates.

	Second stage dependent variables				
	Russian population in 2002	Log income per capita in 2010	Education among indigenous population	Log birth rate in 2011	Share of federal subsidies in local budget in 2007
	(1)	(2)	(3)	(4)	(5)
Russian settlers in 1897, %	1.209*** (0.190)	1.128*** (0.342)	0.354** (0.140)	−1.156*** (0.241)	−0.729*** (0.163)
Russians with degree in 2002, %	0.209 (0.438)	0.578 (0.786)	0.0746 (0.322)	0.0250 (0.553)	−0.529 (0.417)
Log literacy of indigenous population (non-Russian languages)	−0.0214 (0.0469)	0.0333 (0.0843)	−0.0580 (0.0345)	0.0861 (0.0593)	0.0141 (0.0442)
Population density	−0.00349 (0.00281)	−0.00888* (0.00504)	−0.00252 (0.00206)	0.00958** (0.00355)	0.00630** (0.00267)
District area	−0.0307** (0.0113)	−0.0578*** (0.0204)	−0.00858 (0.00835)	0.0250* (0.0143)	0.0334*** (0.0112)
Elevation	−1.04e−05 (0.000133)	8.56e−05 (0.000239)	0.000173* (9.79e−05)	−0.00477*** (0.00168)	−0.000106 (0.000136)
Temperature	0.00781 (0.0144)	0.0429 (0.0259)	0.0281** (0.0106)	−0.0270 (0.0182)	−0.0350** (0.0159)
Second stage R-squared	0.910	0.726	0.274	0.743	0.783
First stage Dependent variable: Russian settlers, %					
Distance to the Black Sea (first stage coefficient)	−0.914*** (0.244)	−0.914*** (0.244)	−0.914*** (0.244)	−0.914*** (0.244)	−0.914*** (0.244)
Full set of controls	Yes	Yes	Yes	Yes	Yes
First stage R-squared	0.833	0.833	0.833	0.833	0.833
First stage F-statistics	15.66	15.66	15.66	15.66	15.66
Observations	30	30	30	30	30

Note: All regressions contain a constant term. Standard errors in parentheses.

* $p < 0.1$.** $p < 0.05$.*** $p < 0.01$.**Table 11**

Long-term effect of middle class. OLS estimates.

	Log income per capita in 2010	Education (% with college or university degree)	Education among Russians	Education among indigenous population	Log birth rate in 2011	Share of federal subsidies in local budget in 2007
	(1)	(2)	(3)	(4)	(5)	(6)
Middle class in 1897	19.90** (8.731)	8.379*** (2.517)	3.682* (2.086)	7.241*** (2.189)	−14.95* (7.731)	−11.03* (6.171)
Log literacy of indigenous population (non-Russian languages)	0.228** (0.0713)	0.0166 (0.0206)	−0.0323* (0.0170)	0.00439 (0.0179)	−0.117* (0.0632)	−0.0841 (0.0504)
Population density	0.00205 (0.00553)	0.000916 (0.00159)	0.000675 (0.00132)	0.000979 (0.00139)	−0.000813 (0.00489)	−0.000960 (0.00349)
District area	0.00701 (0.0195)	0.00982* (0.00562)	−0.00309 (0.00466)	0.0129** (0.00489)	−0.0371** (0.0173)	−0.00623 (0.0141)
Elevation	−0.000133 (0.000201)	2.65e−05 (5.79e−05)	8.64e−05* (4.80e−05)	0.000103* (5.03e−05)	−0.000143 (0.000178)	6.54e−05 (0.000142)
Temperature	0.0171 (0.0330)	−0.000650 (0.00951)	0.00244 (0.00788)	0.0176* (0.00827)	−0.00250 (0.0292)	−0.0163 (0.0221)
Observations	30	30	30	30	30	24
R-squared	0.560	0.457	0.457	0.563	0.346	0.540

Note: All regressions contain a constant term. Standard errors in parentheses.

* $p < 0.1$.** $p < 0.05$.*** $p < 0.01$.

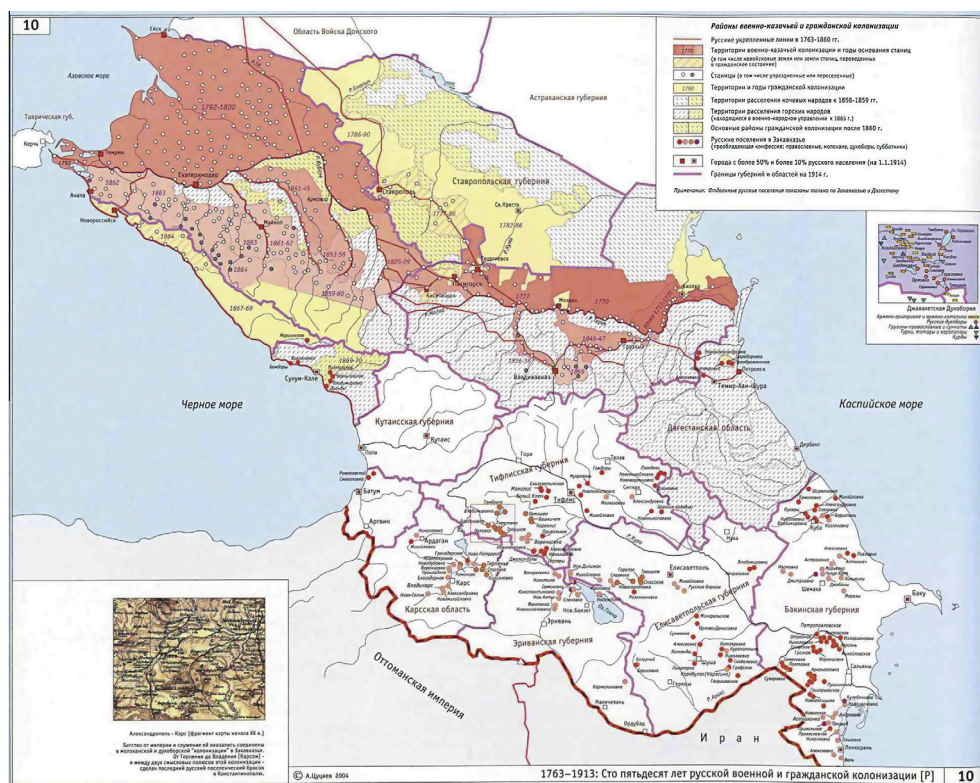


Fig. 15. Russian military and civil colonization of the Caucasus. Source: Tsutsiev (2006). Legend: White dots represent Russian settlements in the North Caucasus. Red and rose dots represent Russian settlements in the South Caucasus. Red and rose areas – areas controlled by the Russian military and Cossacks in the middle of the nineteenth century. Yellow areas – areas of civil colonization by Russian peasants in the late nineteenth century. Dashed (white and yellow) areas – areas of indigenous population in the North Caucasus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Finally, I found strong positive effects of settlement on the current quality of local municipally governance. Today, the North Caucasian region is among the poorest and least-developed regions in the Russian Federation. A significant part of local budgets comes from federal subsidies. I collected data on the share of federal subsidies in districts' budgets from official data sources. Summary statistics in Table 1 indicate that, on average, North Caucasian districts have a high dependency on the federal budget – 73.5% of municipal income comes from federal subsidies. However, there is a high variation from 24% to 97% across the thirty districts. In column (5) of Table 8 I show that colonization history goes a long way in explaining this variation. Districts with higher proportions of Russian population in 1897 have lower shares of federal subsidies in the municipal budget of today (Fig. 14).

Could these correlations be interpreted as the long-term effect of Russian settlement? To answer this question, I again refer to the IV strategy and use distance to the Black Sea coast as an instrument for settlement. The results are reported in Table 10. They completely confirm my previous OLS estimations.

As an additional check in Table 11 I replace the share of Russian settlers with the size of the middle class in 1897 as an explanatory variable and show that this variable predicts income in 2010, educational level in 2002 (share of people with university and college degree) in general, and educational level among both indigenous and Russian population. Note that in case of indigenous people the coefficient is substantially higher and more significant than in case of Russians, indicating that social structure is another potential source of long-term persistence as documented in Acemoglu et al. (2011)

8. Conclusion

This paper documents the importance of human capital and social structure for long-term economic development. I exploit a historical case – the colonization of the North Caucasus region by the Russian Empire – to show that districts with greater proportions of Russian settlers, who were on average more literate than the indigenous people, had higher literacy rates among the indigenous population in the late nineteenth century and today have higher income per capita, higher educational attainment among indigenous populations and higher quality local municipal government.

I use a number of strategies to show that these correlations are not driven by unobserved factors. The qualitative and quantitative evidence show that there was no relationship between Russian settlement and the initial economic devel-

opment level of the districts. The primary purpose of colonization was to protect Russia's southern border and obtain access to warm-water ports on the Black Sea coast. Thus, Russian settlement may be viewed as an exogenous treatment on the indigenous population that varied with proximity to the Black Sea coast. I use distance to the coast as an instrument for the proportion of Russian settlers and confirm my OLS estimates. To insure the validity of the instrument, I conduct a placebo test that shows that distance to the coast is unable to predict literacy in the South Caucasus, an area where Russians had no strategic interest in protecting the coastline. However, Russian settlements predict the literacy of the native population in 1897 and income per capita today in the south, as they do in the north.

I explore several channels of influence – school building policy, social structure and railroads. School construction was more intense in regions with a higher share of Russians and, therefore, benefited the indigenous people who lived closer to Russian settlements in the first place. Additionally, Russian settlers constituted up to 80% of the core middle class occupations (education and health care) that were essential for transmission of human capital. However, railroad construction does not appear to have an effect in terms of variation in subsequent economic development.

There are a few directions for further development of this study. First, it is interesting to explore the effects of Russian colonization on other non-Russian regions of the Empire, such as Central Asia and South Siberia. Second, it is possible to show how the accumulation of human capital affected the cultural norms of local communities. There is qualitative evidence that indigenous people in districts with a higher proportion of Russians today are less traditional, less religious and more inclined to modern secular values. However, without rigorous quantitative analysis, this observation remains speculative.

Third, it is important to explore the political economy of colonization. How were the process of political centralization and the integration of new territories organized? What are the similarities and the differences between Russian colonial experience and the strategies of European powers? What caused these differences? This study is a step towards a better understanding of these issues.

Acknowledgments

I thank Eren Arbatli, Thomas Barret, Pauline Grosjean, Renat Karamoursov, Rinat Menyashev, Charles King, John Nye, Luigi Pascali, Leonid Polishchuk, Ekaterina Zhuravskaya, Maria Yudkevich, participants of 2013 Russian Summer School on Institutional Analysis (RSSIA), participants of 2014 Warwick Summer Workshop in Economic Growth, two anonymous referees and the editor for valuable suggestions and comments. My special thanks to Andrei Markevich for his help with data sources on Russian economic history and Ruben Enikolopov for detailed review that helped to improve the paper.

The financial support of the Center for Fundamental Research at the National Research University Higher School of Economics is greatly appreciated.

References

- Acemoglu, Daron, Johnson, Simon, Robinson, James, 2001. The colonial origins of comparative development: an empirical investigation. *The American Economic Review* 91, 1369–1401.
- Acemoglu, Daron, Johnson, Simon, Robinson, James, 2002. Reversal of fortune: geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics* 117, 1231–1294.
- Acemoglu, Daron, Hassan, Tarik, Robinson, James, 2011. Social structure and development: a legacy of the Holocaust in Russia. *The Quarterly Journal of Economics* 126, 895–946.
- Acemoglu, Daron, Gallego, Francisco, Robinson, James, 2014. Institutions, Human Capital and Development. NBER Working Paper 19933.
- Allen, William Edward David, Muratov, Paul, 1953. *Caucasian Battlefields: A History of the Wars on the Turko-Caucasian Border, 1828–1921*. Cambridge University Press.
- Barrett, Thomas M., 1995. Lines of uncertainty: the frontiers of the North Caucasus. *Slavic Review* 54, 578–601.
- Barrett, Thomas M., 1999. *At the Edge of Empire: The Terek Cossacks and the North Caucasus Frontier, 1700–1860*. Westview Press.
- Berkowitz, Daniel, Caner, Mehmet, Fang, Ying, 2012. The validity of the exclusion restriction revisited. *Journal of Econometrics* 166 (2), 255–267.
- Comrie, Bernard, 2008. Linguistic diversity in the Caucasus. *Annual Review of Anthropology* 37, 131–143.
- Dell, Melissa, 2010. The persistent effects of Peru's mining Mita. *Econometrica* 78 (6), 1863–1903.
- Diamond, Jared, Robinson, James, 2010. *Natural Experiments of History*. Harvard University Press.
- Droller, Federico, 2013. Migration and Long-run Economic Development: Evidence from Settlements in the Pampas. Brown University, Job Market Paper.
- Dunning, Thad, 2012. *Natural Experiments in the Social Sciences. A Design-based Approach*. Cambridge University Press.
- Easterly, William, Levine, Ross, 2012. The European Origins of Economics Development. NBER Working Paper No. 18162.
- Engerman, Stanley, Sokoloff, Kenneth, 1997. Factor endowments, institutions, and differential growth paths among new world economies. In: Haber, Stephen (Ed.), *How Latin America Fell Behind*. Stanford University Press, Palo Alto, CA.
- Engerman, Stanley, Sokoloff, Kenneth, 2002. Factor Endowments, Inequality and Path of Development among New World Economies. NBER Working Paper 9259.
- Gallup, John, Sachs, Jeffrey, Mellinger, Andrew, 1998. Geography and Economic Development. NBER Working Paper 6849.
- Gennaioli, Nicola, La Porta, Rafael, Lopez de Silanes, Florencio, Shleifer, Andrei, 2013. Human capital and regional development. *The Quarterly Journal of Economics*, 105–164.
- Gizetti, A.L., 1901. Statistics on Mortality of the Caucasus Army in Caucasian, Persian and Turkish Wars, 1801–1885. Tiflis, 1901 (Gizetti, A.L. *Sbornik svedenii o poteryakh Kavkazskikh voisk vo vremya voin: Kavkazsko-gorskoj, Persidskikh, Turetskikh i v Zakaspijskom krae, 1801–1885*, Tiflis, 1901).
- Glaeser, Edward L., La Porta, Rafael, Lopez de Silanes, Florencio, Shleifer, Andrei, 2004. Do institutions cause growth? *Journal of Economic Growth* 9, 271–304.
- Gregory, Paul, 1994. *Before Command: The Russian Economy from Emancipation to Stalin*. Princeton University Press.
- Guiso, Luigi, Sapienza, Paola, Zingales, Luigi, 2008. Long-term Persistence. University of Chicago, Mimeo.
- Huillery, Elise, 2009. History matters: the long term impact of colonial public investments in French West Africa. *American Economic Journal – Applied Economics* 1 (2), 176–215.
- Khodarkovsky, Michael, 1999. Of christianity, enlightenment, and colonialism: Russia in the North Caucasus, 1550–1800. *The Journal of Modern History* 71 (1999), 394–430.

- Khodarkovsky, Michael, 2008. North Caucasus during the Russian Conquest, 1600–1850s, Mimeo.
- King, Charles, 2008. *The Ghost of Freedom: A History of the Caucasus*. Oxford University Press, New York.
- La Porta, Rafael, Lopez-de-Silanes, Florencio, Shleifer, Andrei, Vishny, Robert, 1997. Legal determinants of external finance. *Journal of Finance* 52, 1131–1150.
- Nunn, Nathan, 2008. The long-run effects of Africa's slave trades. *The Quarterly Journal of Economics* 123 (1), 139–176.
- Nunn, Nathan, 2009. The importance of history for economic development. *Annual Review of Economics* 1 (1), 65–92.
- Nunn, Nathan, 2014. Historical development. *Handbook of Economic Growth*, vol. 2.
- Nunn, Nathan, Wantchekon, Leonard, 2011. The slave trade and the origins of mistrust in Africa. *The American Economic Review* 101 (7), 3221–3252.
- Rhineland, Hamilton L., 1975. Russia's imperial policy: the administration of the Caucasus in the first half of the nineteenth century. *Canadian Slavonic Papers* 17, 218–235.
- Rhineland, Hamilton L., 1981. The creation of The Caucasian Vicegerency. *Slavonic and East European Review* 59, 15–40.
- Rhineland, Hamilton L., 1990. *Prince Michael Vorontsov: Viceroy to the Tsar*. McGill-Queen's University Press.
- Riquelme, Andres, Berkowitz, Daniel, Caner, Mehmet, 2013. Valid tests when instrumental variables do not satisfy the exclusion restriction. *The Stata Journal* 13 (3), 528–546.
- Rocha, Rudi, Ferraz, Claudio, Soares, Rodrigo R., 2013. Human Capital and Development: The Legacy of European Settlements in Brazil. Mimeo.
- Russo-Turkish Wars, 2013. In *Encyclopedia Britannica*.
- Shonger, R., 1899. Statistics on Educational Institutions in Caucasian School District for Twenty-year Period 1879–1898. Tiflis, 1899 (Shonger R. *Svod statisticheskikh svedeniy po uchebnym zavedeniyam Kavkazskogo uchebnogo okruga za dvadtsatiletniy period 1879–1898*. Tiflis, 1899.).
- Shubina, Elena, Aldashev, Gani, Henry, Sabine, 2014. From Saddlers to Harrows: Agricultural Technology Adoption during Russian Colonization in Kazakhstan. University of Namur, Mimeo.
- Tabellini, Guido, 2008. Institutions and culture. *Journal of the European Economic Association* 6 (2–3), 255–294.
- Tolstoy, Leo, 1863. The Cossacks. *The Russian Messenger* (Russkiy Vestnik), p. 1.
- Tsutsiev, Artur, 2006. *Atlas of Ethno-political History of the Caucasus, 1774–2004*. (Atlas Etnopoliticheskoy istorii Kavkaza, 1771–2004). Moscow, "Europe".
- Wantchekon, Leonard, Klasnja, Marko, Novta, Natalija, 2013. Education and Human Capital Externalities: Evidence from Colonial Benin. Unpublished Manuscript.
- Woodberry, Robert D., 2012. The missionary roots of liberal democracy. *American Political Science Review* 106 (2), 244–274.
- Zhuravskaya, Ekaterina, Grosfeld, Irena, Rodnyansky, Alexander, 2013. Persistent anti-market culture: a legacy of the pale of settlement after the Holocaust. *American Economic Journal: Economic policy* 5 (3), 189–226.