Colonization, Human Capital and Development: The Long-term Effect of Russian Settlement in the North Caucasus, 1890s-2000s

Timur Natkhov[1]

I exploit differences in proportion of Russian settlers in the North Caucasus during colonization to estimate the effect of human capital on long-term development. The main purpose of Russian colonization was to protect its access to the warm-water ports. Therefore settlement was an exogenous treatment on indigenous population that varied depending on the proximity to the Black Sea coast. Instrumenting the share of settlers by the distance to the coast I show their causal positive impact on literacy among the indigenous population in late XIX century with long-term effect on income, educational attainment and quality of local governance today.

Keywords: colonization, human capital, long-run development, natural experiment

JEL Codes: N13, N33, O1, O15

Acknowledgements: I thank Leonid Polishchuk and John Nye for valuable comments that helped to improve the paper.

[1] National Research University Higher School of Economics. Center for Institutional Studies.

E-mail: timur.natkhov@gmail.com

Address: Russian Federation, Moscow, Shabolovka street, 26, office 5308.

INTRODUCTION

Economic studies on the effects of colonization lead to ambiguous conclusions. For instance, Daron Acemoglu, Simon Johnson and James Robinson (2001) show that the effects of colonization vary depending on the institutions that Europeans established in the colony. A small number of settlers due to unfavorable climate and disease environment resulted in extractive institutions that had negative long-term consequences for economic growth. Conversely, William Easterly and Ross Levine (2012) find positive effects on present-day per capita income when even a small minority of European population settles in a region during the colonial period because of the central role played by human capital as a transmission mechanism. Robin Grier (1999) studies a sample of 63 former colonies and finds a positive and significant correlation between colonial duration and a country's economic growth during the 1961-1990 period. He also looks for causal channels and concludes that investment in education explains the positive effects of colonial duration. Similarly, James Feyrer and Bruce Sacerdote (2009) find a robust positive relationship between the number of years spent as a European colony and current GDP per capita on a sample of Atlantic, Pacific and Indian Ocean islands. On the most general level, Arhan Ertan and Lous Putterman (2012) find that neither colonization nor its timing have direct effects on income, although they find that colonization – and its timing – may exert an indirect impact on income by affecting the quality of institutions and the European-descended portion of the population. This result was recently confirmed on the regional level by Federico Droller (2013), who finds that areas of Argentina with historically higher proportions of European population currently have significantly higher per-capita GDP and educational levels.

Although these studies establish a persuasive causal link between colonization and development, they suffer from two limitations. First, these studies do not distinguish between the effects of colonization on the overall population and the effects of colonization on indigenous

populations. Second, they rely mostly on cross-country (rarely cross-regional) comparisons that are far from the "ceteris paribus" condition even when controlling for a full set of observable variables.

In this paper, I present a historical case – the colonization of the North Caucasus region by the Russian Empire in the late XIX century – that provides an opportunity to fill this gap for two reasons. First, it allows for the empirical investigation of the effect of settlement on *local* populations and subsequent regional development. Second, it copes with unobserved heterogeneity better than a cross-country study because of the small size of the region and because of common history.

Based on a unique dataset of thirty North Caucasian districts collected from various historical records and the 1897 census of the Russian Empire, I argue that Russian settlement resulted in human capital accumulation among the native population. In a preferred specification, a 10% increase in the Russian-speaking population is associated with an increase in the literacy rate of more than 4% among the indigenous population. Moreover, this effect had long-term consequences for the educational attainment of the local population today, in addition to affecting income levels and the quality of local municipal management in the districts.

Although these correlations are robust to various controls, it cannot be taken as conclusive evidence that the Russian settlement *caused* human capital accumulation and long-run development. An alternative explanation is that Russians choose to settle in districts with higher initial literacy and higher initial levels of development. In other words, the settlement may be correlated with unobserved district characteristics that lead to biased estimates of the settlement effect. Using statistical and historical evidence, I show that this is not the case. Both sources of evidence indicate that the selection of settlement areas was unrelated to their initial levels of development.

First, on a smaller sample of districts for which data are available, I show that there is no correlation between initial population density and subsequent Russian settlement. Similarly, using census data, I show that there is no correlation between the initial literacy rate of the local population

in languages other than Russian and the share of Russian settlers. Thus, Russians did not select themselves into settling in more literate and developed districts.

I also review the historical literature on the reasons for North Caucasus colonization; this confirms my statistical results. This review shows that the main purpose of colonization for Russian Empire was to protect its southern frontier from its constant rival – the Ottoman Empire – and to establish and protect warm-water ports on the Black Sea to expand trade routes into the Mediterranean region. Thus 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. This resulted in a substantial cross-district variation in the share of Russian settlers, which was exogenous to local population.

Second, I use distance to the Black Sea as an instrument to estimate the causal effect of Russian settlement on indigenous populations and subsequent economic development. Similar to OLS coefficients, the IV coefficients are positive and significant, which suggests that increased proportion of Russians settlers in a district positively affected literacy rates among the indigenous populations in the late nineteenth century.

Moreover, I show that this 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 and higher-quality municipal management. IV results again confirm OLS estimations, suggesting that Russian settlement had a causal long-term effect on development of North Caucasian districts.

In contrast to other studies of colonization, the historical case exploited by this paper focuses on one relatively small region with common history, institutions and cultural norms. This has an advantage because it is much closer to the ceteris paribus condition than a cross-country study or even a cross-regional study within one country. I show that holding institutions and culture fixed it is human capital that makes a difference and drives long-term development.

The remainder of this paper is structured as follows. In the next section, I provide a short historical overview of the North Caucasus region and the causes and manner of colonization by the Russian Empire. In Section III, I describe the data used to estimate the effects of colonization. Section IV documents the correlations in the data, and Section V addresses causality. In Section VI, I examine the long-term consequences of Russian settlement. Section VII concludes.

HISTORICAL BACKGROUND

The North Caucasus is the northern part of the Caucasus region between the Black and Caspian Seas and lies within European Russia. The region is characterized by enormous ethnic and language heterogeneity. The relatively small territory contains 46 different languages in 8 language families.

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, as the ones south of the Caucasian range in Georgia, Armenia and Azerbaijan. Located on the peripheries of the Ottoman, Persian, and Russian 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 northeast Caucasus. By the mid-nineteenth century, however, the Russians succeeded in turning the Caucasus from a contested frontier zone into the borderland of the Russian Empire.

After the victory in the Russo-Turkish War of 1768-1774, the Russian Empire gained Kabarda in the central part of North Caucasus, the port of Azov (in the North West) and other ports of the Crimean peninsula as a result of the Treaty of Kuchuk Kaynarca. These new territories allowed Russia to move its southern frontier closer to the lands of local people. Beginning in the late eighteenth century, the Empire began its expansion into 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).

The primary method of expansion was the construction of a chain of forts and garrisons along the Terek, Malka and Kuban rivers. In 1735, the Russian government built the Kizlyar fortress, which is located in the Terek River delta. This laid the foundation for the chain of forts later known as the Caucasian line (*Kavkazskaya liniya*). In the 1770s, almost half of the Volga Cossacks were resettled around Mozdok – another fortress on the Terek River. Other major fort posts for Russian expansion into the Caucasus were founded by settlers, including Giorgiyevsk and Stavropol in 1777, and Vladikavkaz in 1784. By 1825, a continuous chain of fortifications stretching from the Caspian to the Black Sea firmly separated the plains – overwhelmed by Russian settlers – from the foothills and mountains in which the native inhabitants continued to reside (see Map 1). As military and political expansion continued throughout the nineteenth century, the frontier gradually shifted south and included more lands with indigenous populations. Thus, the southern Russian frontier became a moving boundary and the contact area between the Russian and the North Caucasian societies.

Russian colonization strategies ranged from violent military conquest to co-opting local elites; Russia dispatched bureaucrats, merchants, and priests to the newly built towns, and encouraged newcomers from neighboring Russian regions to settle in and farm the land. The latter method has become the main colonization path since the late nineteenth century, when major military operations were completed. Russia's strategic goals were to integrate the region and its peoples into the Russian Empire (Khodarkovsky, 1999).

However, in practice, such integration meant a thorough transformation of local and tribal identities into modern, imperial, and Russian identities; this was not easy to achieve. The agents of such transformation were Slavic settlers from the interior and the Cossacks – a group of Slavic independent semi-military communities. The historical evidence suggests that the arrival of new settlers in the North Caucasus region encouraged local populations to trade and engage in other close

relationships with Russian settlers. Thus, the indigenous people learned to speak and write in Russian. Because most of the local languages were unwritten, the Russian language formed the basis for local alphabets. Today Russian serves as the region's lingua franca and is spoken by nearly everyone.

DATA

To explore the effects that Russian settlement had on indigenous population, I construct a dataset using a number of sources. The first dataset comes from the Russian Imperial Census conducted in 1897, the only nationwide census conducted in the Russian Empire. It 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 Caucasus.[2] These Caucasian regions are subdivided into thirty smaller districts that represent my main sample.

The data from the Russian Imperial Census of 1897 allows people to be distinguished by their ethnicity (defined in terms of native language). For each of the thirty districts, the census provides data on total population, numbers 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 proportions of Russians in the districts and literacy rates among both Russians and the indigenous population. I match this historical data with modern data sources on income and quality of municipal management from Rosstat (the Russian Official Statistical Agency), educational attainment and birth rates from the latest available Population Census 2002 and with various geographical characteristics, such as temperature, elevation, etc.

Although the number of observations (thirty districts) seems small for statistical analysis, this database has one advantage. Table 1 reports summary statistics and shows enormous variation in

[2] These are Chernomorskaya, Kubanskaya, Stavropolskaya, Terskaya and Dagestanskaya oblasts.

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%, and the literacy among Russian settlers – from 11.5% to 78.6%, population density – from 2.6 to 55.8 people on a square kilometer. As the following empirical analysis will show, this variation provides a good opportunity to estimate the effects of settlement on regional development in a long-term perspective.

BASIC CORRELATIONS: 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:

$$Ln\ Literacy\ Indig_i = \beta_0 + \beta_1 Russians\ Share_i + \beta_2 Russians\ Literacy_i + \gamma \boldsymbol{X}_i' + \varepsilon_i \quad (1),$$

where $Ln\ Literacy_i$ is the natural log of the literacy rate among indigenous population in the i district, $Russians\ Share_i$ is the share of Russian settlers in the district, $Russians\ Literacy_i$ is the literacy rate among Russian settlers and $\boldsymbol{X}_i'$ is a vector of control variables that capture differences in districts' economic development, geography and climate.

Table 2 reports the estimation results for equation (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 indigenous people's literacy is positive and statistically significant. In columns (3) to (6), 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.

One concern with the estimates in columns (1) through (6) is that they may be biased because of the inclusion of districts with small numbers of settlers or small numbers of indigenous people. According to the census data from 1897, there is one district with fewer than fifty Russian settlers (in southern Dagestan) and two districts with less than fifty indigenous persons (one in Kuban and one in the Stavropol region). Dropping these tree observations in column (7) does not present a difference.[3]

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 between 0.91 to 1.05 standard deviations in literacy rates 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 4% in the indigenous population. Figure 1 shows the estimation results from column (7).

CAUSALITY ISSUES: DETERMINANTS OF SETTLEMENT AND IV APPROACH

The OLS estimates show that there is a strong and significant relationship between Russian settlement and literacy among the indigenous population in 1897. However, it remains unclear whether Russian-speaking settlers have a causal impact on the local population in this regard. An alternative explanation for this relationship is that Russians selected districts for settlement with higher initial literacy and higher development levels in general, and that, therefore, we observe a

[3] Dropping even five observations (out of a thirty-district sample) with less than one-hundred settlers and fewer than one-hundred indigenous people makes almost no difference both in terms of statistical significance and the magnitude of the coefficient of interest.

positive correlation between the proportion of settlers and the literacy of the local population. In other words, this estimation suffers from omitted variable bias.

I this section, I pursue two strategies to evaluate whether there is a causal effect of settlement on literacy. First, using historical evidence and statistical data, I explore the most important determinants of settlement. As I will show, these sources suggest that settlers' self-selection was important, but it was unrelated to the initial literacy rate or initial development level of local population. The main purpose of colonizing the North Caucasus region for the Russian Empire was to protect its southern frontier against the Ottoman Empire and to obtain access to the Black Sea. Therefore, settlement strategy followed a strong geographical pattern – the western districts of North Caucasian region (closer to the east coast of the Black Sea) were settled by Russians more extensively than eastern districts. This resulted in a geographical variation of the proportion of Russian settlers across districts.

A simple correlation between distance to the Black Sea and the proportion of Russian settlers in the district is high (-0.82) and statistically significant at a 1% level. As a second strategy, I use distance to the coast as an instrument for Russian settlement. The IV (instrumental variable) results confirm my OLS estimates.

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? I collected data on the literacy rate among indigenous populations in languages other than Russian to answer these questions. This is a good proxy for the initial literacy of indigenous people for the following reasons. Before Russian colonization, the indigenous languages were unwritten. Historically, two languages served as lingua franca – Turkish and Arabic. The first was used mostly by merchants in

the western and central regions of the North Caucasus, whereas the second was more common in Dagestan (eastern part) and used by religious authorities because the majority of the population were Muslim. The census data from 1897 show that the districts with the highest initial literacy rates were in Dagestan (10-15% of adult population was literate) and districts with the lowest literacy rates were in the central part of the North Caucasus (Ossetia and Kabarda). Next, I collect the earliest available data on population density in the district.[4] Before massive colonization started in the late 1860s, Russian government collected preliminary statistical information about the region.[5] These data are available only for 21 districts in my sample.

Both initial literacy and initial population density serve as proxies for the initial development levels 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 initial literacy rate is positive but statistically insignificant – there is no relationship between Russian's decision to settle and the initial literacy of indigenous population.

Column (2) uses 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 Russians selected themselves into regions with lower population densities and, therefore, a lower level of development, but this effect disappears when I extensively control for several geographical

[4] 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, Johnson, Robinson (2002)

[5] 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 tablitza Kavkazskogo kraya v 1865"*. Tiflis, 1866)

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 (Figure 2). This empirical finding is consistent with established historians' view of North Caucasus colonization. For instance, Rhinelanders 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, 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" (Rhinelanders, 1975, p. 218).

To understand the Russian Empire's strategic colonization of the North Caucasus region, it is important to place this event into a broader historical context. Historical evidence shows that geopolitical and military reasons were the most important determinants of the nature of settlement. For instance, Khodarkovsky notes 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 main rival for these coveted regions for many years was the Ottoman Empire. Russian-Turkish wars extended into some of the longest conflicts in human history. For two centuries, from 1700 to 1900, there were eight significant military conflicts between Russia and Turkey, every 25 years on average. The early wars were mostly sparked by Russia's attempt to establish a warm-water port on the Black Sea. Subsequent wars were fought to gain control of the Dardanelles and the Bosporus straits, to retain access to world trade routes and to expand into the Caucasus. Russia's victory in the 1828-1829 war resulted in the Treaty of Edirne that gave Russia most of the eastern shore of the Black Sea (Encyclopedia Britannica, 2013). Thirty years later, after the emancipation of Russian serfs in 1861, the Russian government began an active policy of settling this region with peasants from neighboring regions in addition to Cossacks and the military (see Map 1).

All these facts indicate that the eastern side of the Black sea coast was the strategic goal of North Caucasus colonization. In fact, 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 twenty-one 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 initial literacy and a number of geographical variables such as district area, elevation and temperature, the only variable that is significantly (statistically and economically) predicts Russian settlement is the distance to the Black Sea coast. Figure 3 shows this relationship as estimated in column (5).

Instrumental Variable

Both historical and statistical evidence suggests that Russian settlement was unrelated to the characteristics of the local population. Thus, we can see settlement as "as if random" treatment, which generate the natural experiment of history (Diamond, Robinson, 2010; Dunning, 2012). I can 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 equation (1).

The validity of the instrument relies on two properties. First it must be correlated with Russian settlement, and second it must not be correlated with other characteristics that might affect the literacy of indigenous population. While the first property is confirmed in Table 3, the second could not be tested explicitly. However, I can test whether distance to the Black Sea is correlated with initial development by using data on initial literacy and initial population density. As the data show, the correlation coefficients between initial literacy or initial population density and distance to the Black Sea are small and statistically insignificant (-0.26 and 0.21 respectively). On the contrary, the correlation between Russian language literacy and distance to the Black Sea is high (-0.63) and significant at a 1% level. Thus distance to the Black Sea coast was unrelated to initial development and became important for local population only *after* the Russians began to settle. In other words, distance to the coast affected the literacy of the indigenous population only through Russian settlement. I see this as persuasive evidence for the exclusion restriction to be valid.

The results of the IV estimation are reported in Table 4. The first column reports estimates without control variables, the second includes the literacy of settlers, and the third includes development and geography controls. In column (4), I exclude districts with fewer than fifty Russian settlers and fewer than fifty persons representing indigenous people. The first stage estimates 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 the number of Russians who settled there. The first stage F-statistics are also high in all specifications.

Overall, the IV results confirm the OLS estimates. The positive relationship between Russian settlement and the literacy of the indigenous population in the late nineteenth century reflects the causal effect of settlement.

THE LONG-RUN EFFECT OF RUSSIAN SETTLEMENT

Finally, 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 5, column (1), I take into account the current educational level of the Russian population and the same set of development and geography 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 (Figure 4).

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 (see Figure 5). This is the effect on the entire population of the district because, unfortunately, the data do not permit 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 (Figure 6). 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 that finding, 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 (see column (4) and Figure 7).

Finally, I found strong positive effects of settlement on the current quality of local municipally government. 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), I show that Russian settlement 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, which indicates higher-quality local government (Figure 8).

Could these correlations be interpreted as the causal 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 6.[6] They completely confirm my previous OLS estimations.

[6] I use the same set of dependent and independent variables as in Table 5.

CONCLUSION

This paper documents the importance of human capital 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. This effect is persistent – districts that had higher proportions of Russian settlers in the late nineteenth century continue today to have higher income per capita, higher educational attainment among indigenous populations, lower birth rates and higher quality local municipal government.

I pursued a number of strategies to show that this correlation is not driven by self-selection of settlers and reflects the causal effect of Russian settlement. The qualitative and quantitative evidence show that there was no relationship between Russian settlement and the initial literacy of the indigenous people or the initial economic developmental level of districts. Historians' works suggest that colonizers "cared little for the peculiarities of Caucasian life," and the main purpose of colonization was to protect Russia's southern border with Turkey and access to the warm-water ports on the Black Sea coast. Thus, this case provides a natural historical experiment in which Russian settlement may be viewed as an exogenous treatment on the indigenous population that varied with the proximity to the Black Sea coast. Moreover, distance to the Black Sea coast does not correlate with the initial development level of the districts affecting the literacy rates of indigenous people only through Russian settlement. Thus, the exclusion restriction holds and I can use distance to the coast as an instrument for the proportion of Russian settlers. The IV estimates confirmed the OLS results, suggesting that when more Russians settled in the district, the literacy rate of the indigenous people in the late nineteenth century rose; higher income per capita and other development indicators remain in such districts today.

Perhaps the best evidence to support of this conclusion is the opinion of the people themselves. Photo 1 shows "The Monument to a Russian Teacher" that was built in 2006 in Makhachkala – the capital of the most populous North Caucasian region. It pays tribute to Russian teachers who dedicated their lives to teaching the Russian language to the Dagestan peoples. Similar monuments and memorial plaques with the names of teachers still stand today in many smaller villages and communities in the North Caucasian region.

REFERENCES

Acemoglu, Daron, Simon Johnson and James Robinson, "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *The Quarterly Journal f Economics,* 117 (2002), 1231–1294.

Acemoglu, Daron, Simon Johnson and James Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation." *The American Economic Review,* 91 (2001), 1369–1401.

Diamond, Jared, and James Robinson, *Natural Experiments of History,* (2010). Harvard University Press.

Droller, Federico, "Migration and Long-Run Economic Development: Evidence from Settlements in the Pampas", Brown University, Job Market Paper (2013).

Dunning, Thad, *Natural Experiments in the Social Sciences. A Design-Based Approach,* (2012). Cambridge University Press.

Easterly, William, and Ross Levine, "The European Origins of Economics Development." NBER Working Paper No. 18162, Cambridge, MA, June 2012.

Ertan, Arhan, Lous Putterman, and Martin Fiszbein, "Determinants and Economic Consequences of Colonization: A Global Analysis." Mimeo, (2012).

Feyrer, James and Bruce Sacerdote, "Colonialism and Modern Income: Islands as Natural Experiments." *The Review of Economics and Statistics*, 91 (2009), 245–262.

Grier, Robin, "Colonial Legacies and Economic Growth." *Public Choice* 98 (1999), 317-334

Khodarkovsky, Michael, "North Caucasus during the Russian Conquest, 1600-1850s." Mimeo, (2008).

Khodarkovsky, Michael, "Of Christianity, Enlightenment, and Colonialism: Russia in the North Caucasus, 1550–1800." *The Journal of Modern History*, 71 (1999), 394-430.

Rhinelanders, Hamilton L., "Russia's Imperial Policy: The Administration of the Caucasus in the First Half of the Nineteenth Century." *Canadian Slavonic Papers*, 17 (1975), 218-235.

Russo-Turkish wars, (2013). In *Encyclopædia Britannica*.

Tsutsiev, Artur, "Atlas of Ethno-Political History of the Caucasus, 1774-2004." (*Atlas Etnopoliticheskoy istorii Kavkaza, 1771-2004*). Moscow, "Europa" (2006).

Table 1. Summary Statistics.

	Mean	Std. deviation	Min	Max	N
Historical data (1897 census and 1860s statistics)					
Literacy in Russian among indigenous population, %	5.23	8.33	0.10	32.00	30
Literacy in Non-Russian languages among indigenous population, %	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
Modern Data (2002 Census and Rosstat Data)					
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
Geography					
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	1 658.00	30
Average temperature in January, Celsius	2.32	2.54	-1.34	7.97	30
Distance to the Black Sea, km	670.0	368.1	51.0	1 265.0	30

Table 2. Russian Settlers and Literacy among Indigenous Population in 1897.

	Dependent variable: *Log literacy rate among indigenous population*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Russian settlers	2.773***	3.547***	3.539***	4.007***	4.103***	4.012***	4.646***
	(0.646)	(1.010)	(1.039)	(1.116)	(1.172)	(1.213)	(1.099)
Settlers' literacy rate		2.377	2.361	1.644	1.325	1.053	1.290
		(2.383)	(2.449)	(2.521)	(2.735)	(2.860)	(2.948)
Population density			0.00128	-0.00951	-0.0110	-0.00873	-0.0164
			(0.0247)	(0.0264)	(0.0273)	(0.0283)	(0.0246)
District area				-0.0968	-0.0962	-0.0783	-0.143
				(0.0867)	(0.0883)	(0.0998)	(0.0911)
Elevation					0.273	0.536	-0.0655
					(0.809)	(1.039)	(1.045)
Temperature						0.0629	-0.123
						(0.152)	(0.144)
Constant	-5.665***	-6.733***	-6.753***	-5.846***	-5.899***	-6.22***	-5.125***
	(0.401)	(1.144)	(1.229)	(1.469)	(1.504)	(1.717)	(1.617)
Observations	30	30	30	30	30	30	27
R-squared	0.397	0.419	0.419	0.446	0.449	0.453	0.587

Note: Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 3. Determinants of Russian Settlement.

	Dependent variable: *Share of Russian settlers in 1897*				
	(1)	(2)	(3)	(4)	(5)
Log initial literacy	0.107				0.0217
	(0.0772)				(0.0418)
Log initial population density		-0.223**	-0.111		
		(0.102)	(0.0656)		
Elevation			0.1240	-0.0766	-0.0927
			(0.1053)	(0.119)	(0.124)
Temperature			0.0456*	-0.0354*	-0.0332
			(0.0231)	(0.0191)	(0.0198)
District area			0.0190	0.0246**	0.0252**
			(0.0154)	(0.0105)	(0.0107)
Distance to the Black Sea			-0.754***	-0.853***	-0.814***
			(0.143)	(0.136)	(0.157)
Constant	0.798***	0.903***	0.911**	0.991***	1.030***
	(0.247)	(0.266)	(0.346)	(0.178)	(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<0.01, ** p<0.05, * p<0.1.

Table 4. Russian Settlers and Literacy among Indigenous Population in 1897 (IV estimation).

	Second stage			
	Dependent variable: *Log literacy rate among indigenous population*			
	(1)	(2)	(3)	(4)
Russian settlers, %	3.344***	5.587***	5.644***	6.794***
	(0.796)	(1.777)	(1.831)	(1.836)
Settlers' literacy rate		6.080	2.905	4.258
		(3.616)	(3.331)	(3.746)
Population density			-0.0206	-0.0304
			(0.0309)	(0.0284)
District area			-0.132	-0.202*
			(0.112)	(0.107)
Elevation			0.645	-0.221
			(1.083)	(1.145)
Temperature			0.0257	-0.194
			(0.161)	(0.163)
Constant	-5.934***	-8.789***	-6.868***	-6.037***
	(0.459)	(1.877)	(1.859)	(1.860)
Second stage *R*-squared	0.380	0.331	0.410	0.508
	First stage			
	Dependent variable: *Russian settlers, %*			
	(1)	(2)	(3)	(4)
Distance to the Black Sea	-0.921***	-0.626***	-0.784***	-0.756***
	(0.120)	(0.157)	(0.172)	(0.196)
Settlers' literacy rate		-0.870**	-0.0945	-0.155
		(0.330)	(0.388)	(0.502)
Population density			0.00482	0.00485
			(0.00339)	(0.00365)
District area			0.0291**	0.0288**
			(0.0115)	(0.0132)
Elevation			-0.0582	-0.0311
			(0.129)	(0.163)
Temperature			-0.0265	-0.0207
			(0.0216)	(0.0255)
Constant	1.089***	1.148***	0.804***	0.774***
	(0.0915)	(0.0861)	(0.224)	(0.255)
Observations	30	30	30	27
R-squared	0.677	0.743	0.835	0.809
First Stage *F*-statistics	58.70	39.05	19.41	14.11

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

Table 5. Persistent Effect of Russian Settlers. OLS estimates.

	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, %	1.005***	1.011***	0.0933*	-0.865***	-0.573***
	(0.0890)	(0.174)	(0.0456)	(0.118)	(0.103)
Russians with degree in 2002, %	0.268	0.429	0.727***	-0.537	-0.678*
	(0.319)	(0.624)	(0.164)	(0.423)	(0.328)
Population density	-0.00167	-0.00701	-0.000488	0.00764**	0.00494*
	(0.00236)	(0.00461)	(0.00121)	(0.00312)	(0.00237)
District area	-0.0199**	-0.0478**	0.00761	0.00989	0.0199*
	(0.00879)	(0.0172)	(0.00450)	(0.0116)	(0.0103)
Elevation	-0.0842	0.0715	0.0254	-0.310**	-0.0802
	(0.0834)	(0.163)	(0.0427)	(0.110)	(0.0948)
Temperature	0.00957	0.0522**	0.0245***	-0.0210	-0.0370**
	(0.0124)	(0.0242)	(0.00634)	(0.0164)	(0.0137)
Constant	0.0483	2.065***	-0.188**	3.258***	1.142***
	(0.171)	(0.335)	(0.0877)	(0.227)	(0.190)
Observations	29	29	29	29	29
R-squared	0.923	0.733	0.709	0.767	0.810

Note: Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Persistent 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, %	1.176***	1.221***	0.206***	-0.997***	-0.677***
	(0.125)	(0.234)	(0.0672)	(0.158)	(0.123)
Russians with degree in 2002, %	0.219	0.368	0.695***	-0.499	-0.628*
	(0.346)	(0.645)	(0.185)	(0.436)	(0.340)
Population density	-0.00329	-0.00901*	-0.00156	0.00889**	0.00590**
	(0.00266)	(0.00497)	(0.00143)	(0.00335)	(0.00251)
District area	-0.0285**	-0.0585***	0.00189	0.0166	0.0249**
	(0.0103)	(0.0193)	(0.00554)	(0.0130)	(0.0110)
Elevation	-0.0261	0.143	0.0637	-0.354***	-0.128
	(0.0942)	(0.176)	(0.0505)	(0.119)	(0.102)
Temperature	0.00754	0.0497*	0.0232***	-0.0195	-0.0382**
	(0.0134)	(0.0250)	(0.00719)	(0.0169)	(0.0142)
Constant	0.0594	2.079***	-0.180*	3.249***	1.147***
	(0.185)	(0.346)	(0.0992)	(0.233)	(0.196)
Second Stage *R*-squared	0.910	0.715	0.628	0.754	0.798
Distance to the Black Sea (First Stage coefficient)	-0.810***	-0.810***	-0.810***	-0.810***	-0.810***
	(0.144)	(0.144)	(0.144)	(0.144)	(0.144)
Full set of controls	yes	yes	yes	yes	yes
First Stage *R*-squared	0.826	0.826	0.826	0.826	0.826
First Stage *F*-statistics	17.50	17.50	17.50	17.50	17.50
Observations	29	29	29	29	29

Note: Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

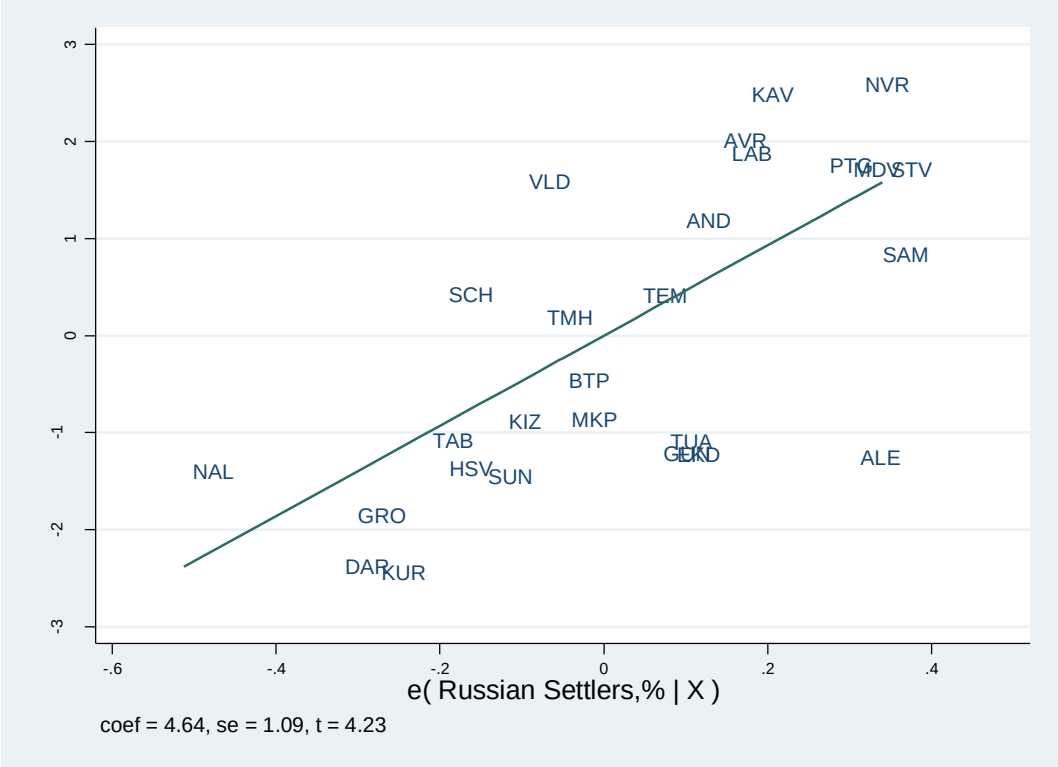


Figure 1. Russian settlers and literacy among indigenous population in 1897.

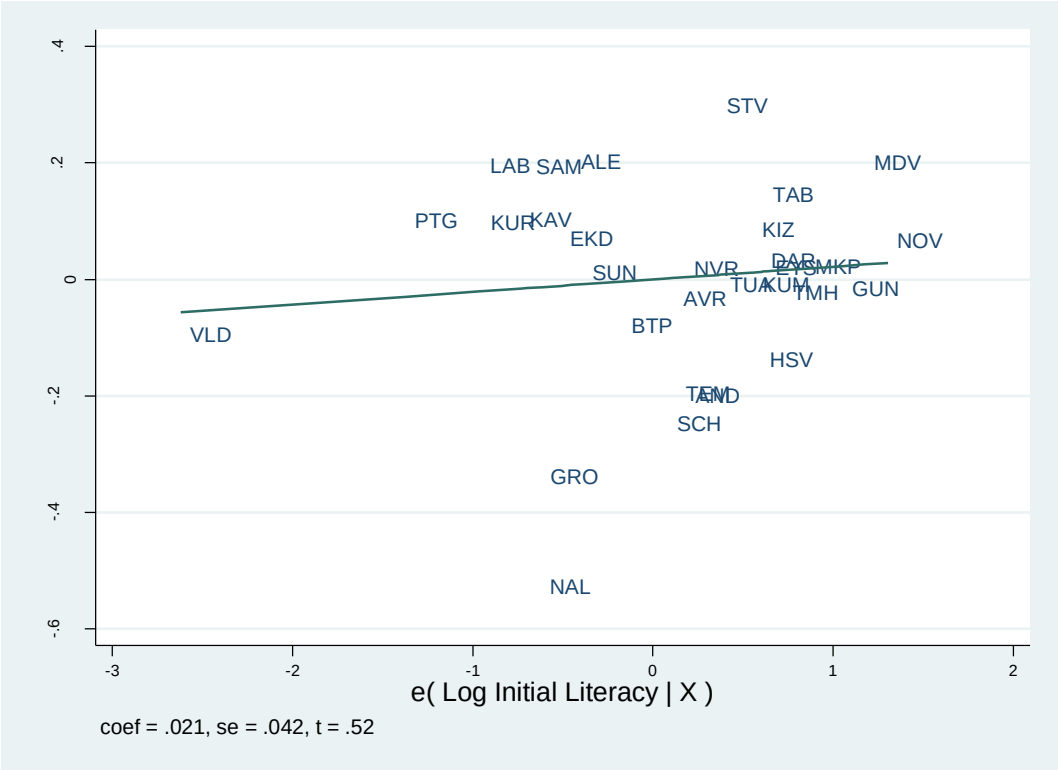


Figure 2. Initial literacy among indigenous population and Russian settlement.

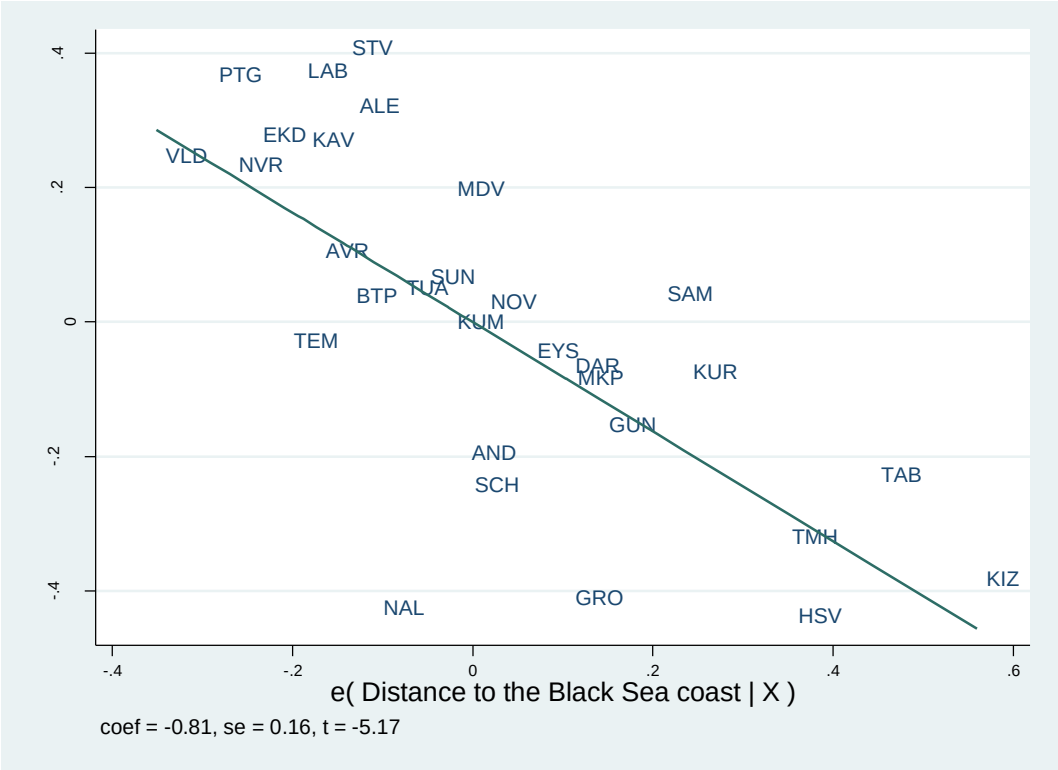


Figure 3. Distance to the Black Sea and Russian Settlement.

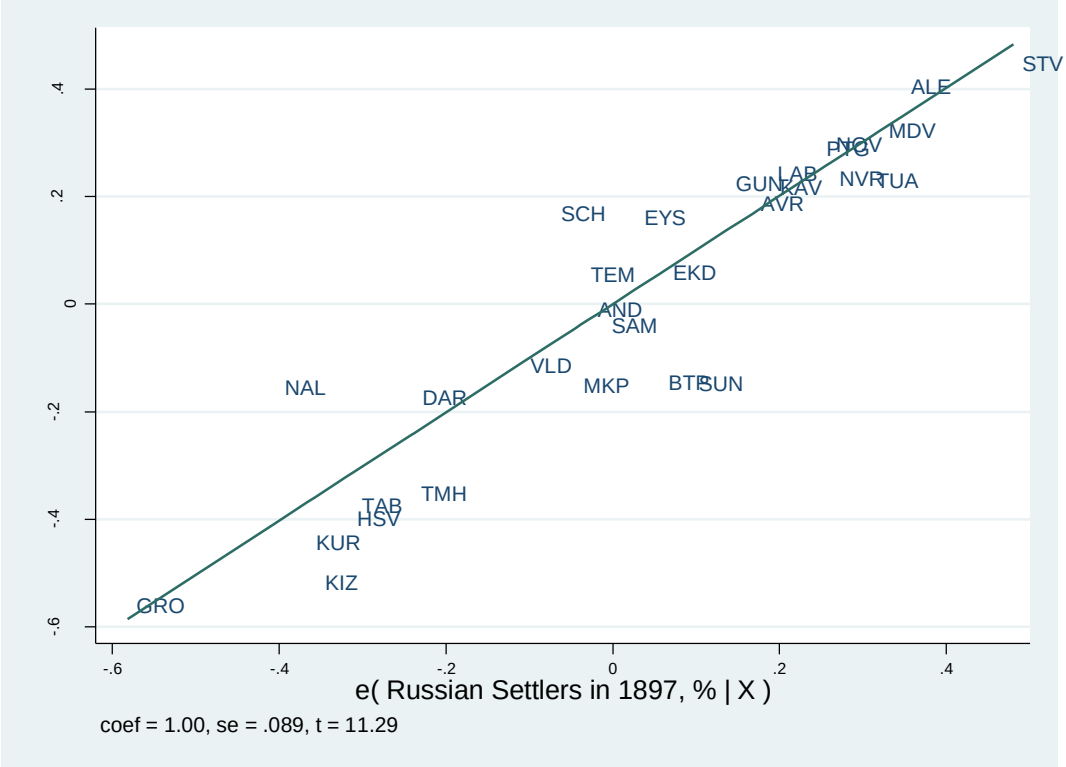


Figure 4. Russian settlement in 1897 and proportion of Russians in 2002.

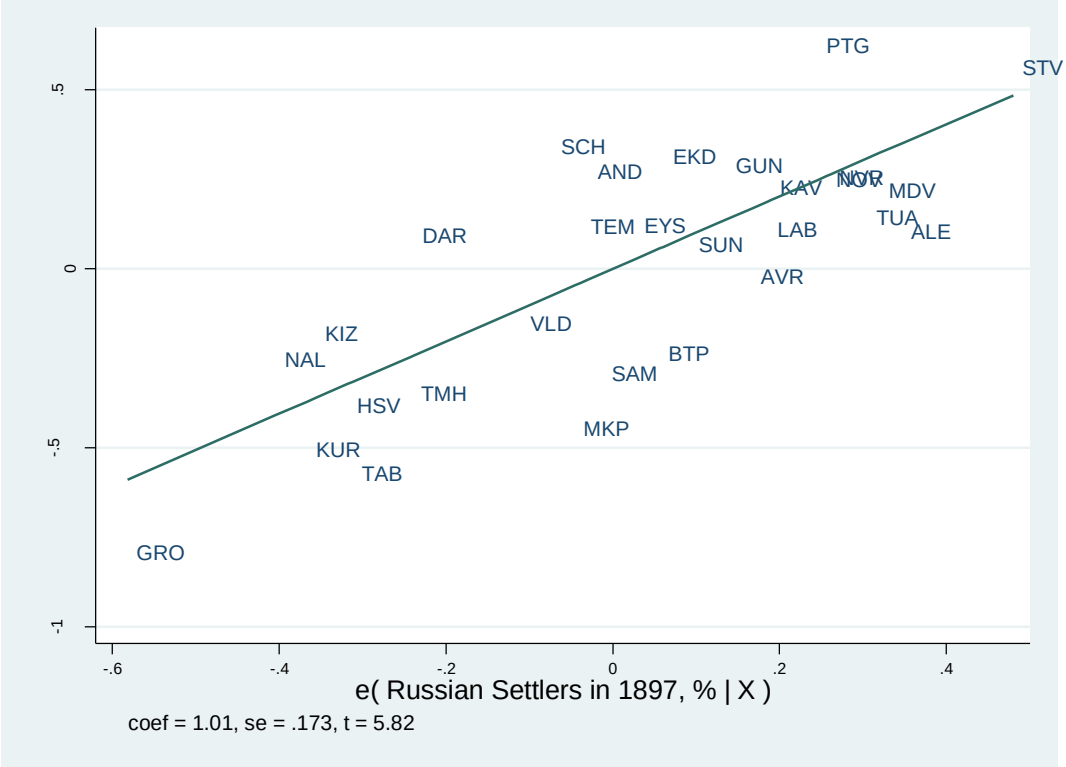


Figure 5. Russian settlement in 1897 and income per capita in 2010.

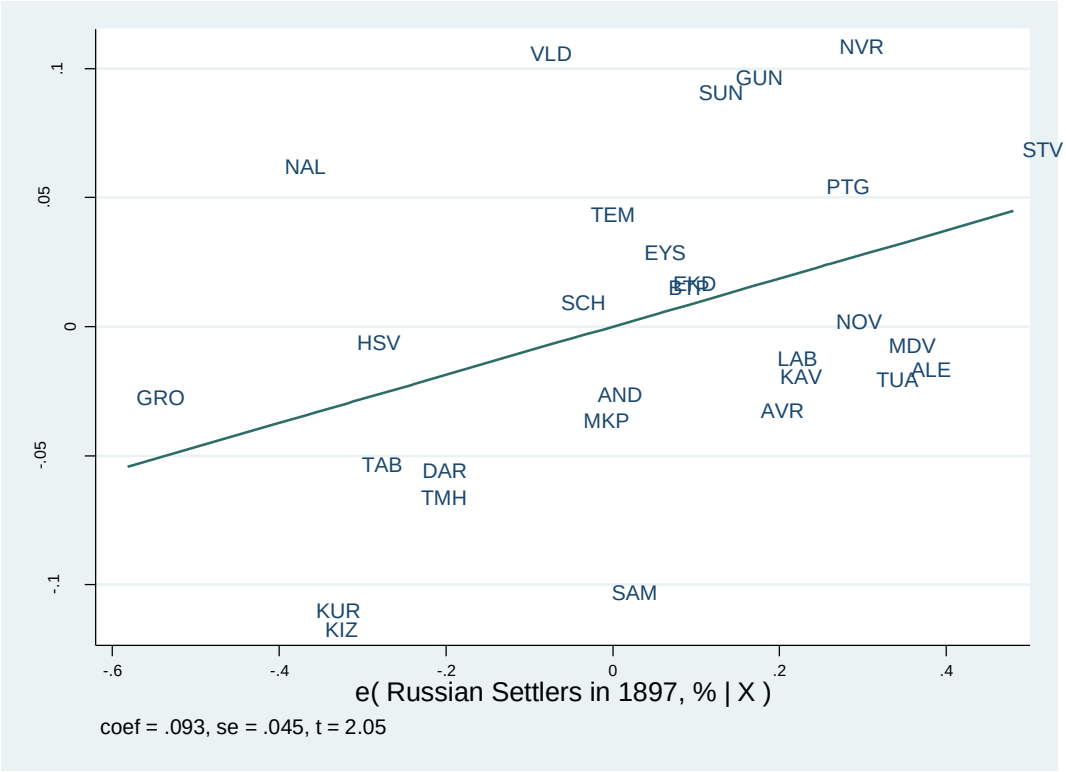


Figure 6. Russian settlement in 1897 and indigenous population with college degree in 2002.

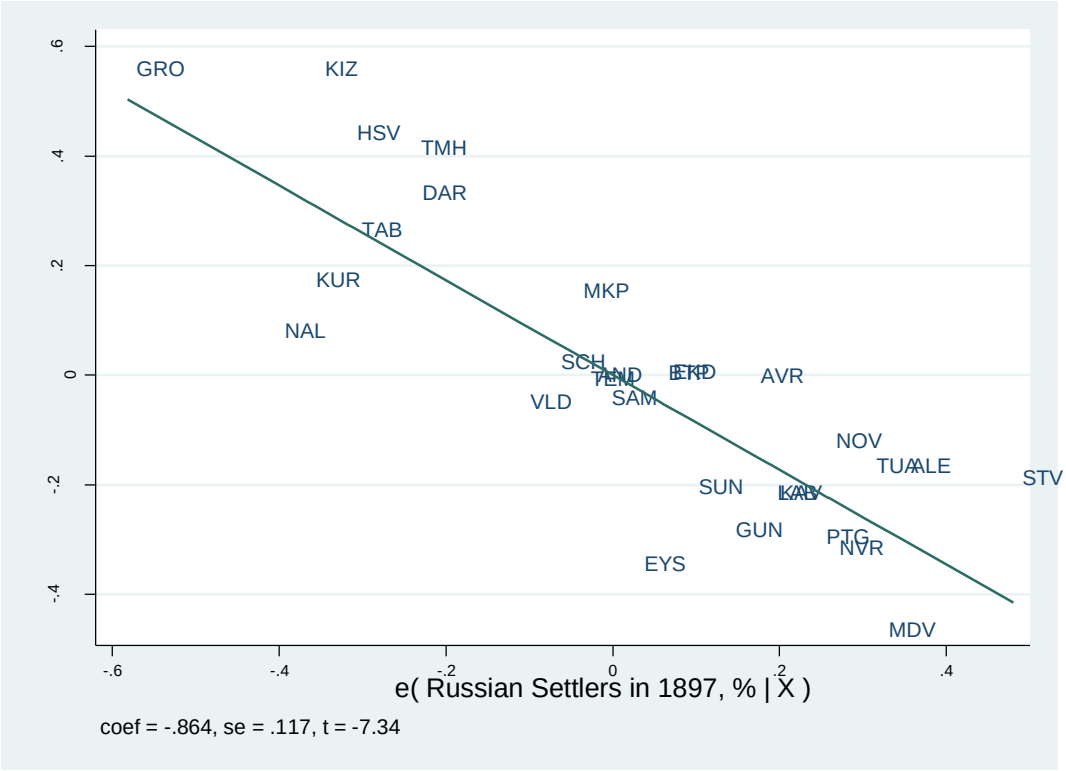


Figure 7. Russian settlement in 1897 and birth rates in 2011.

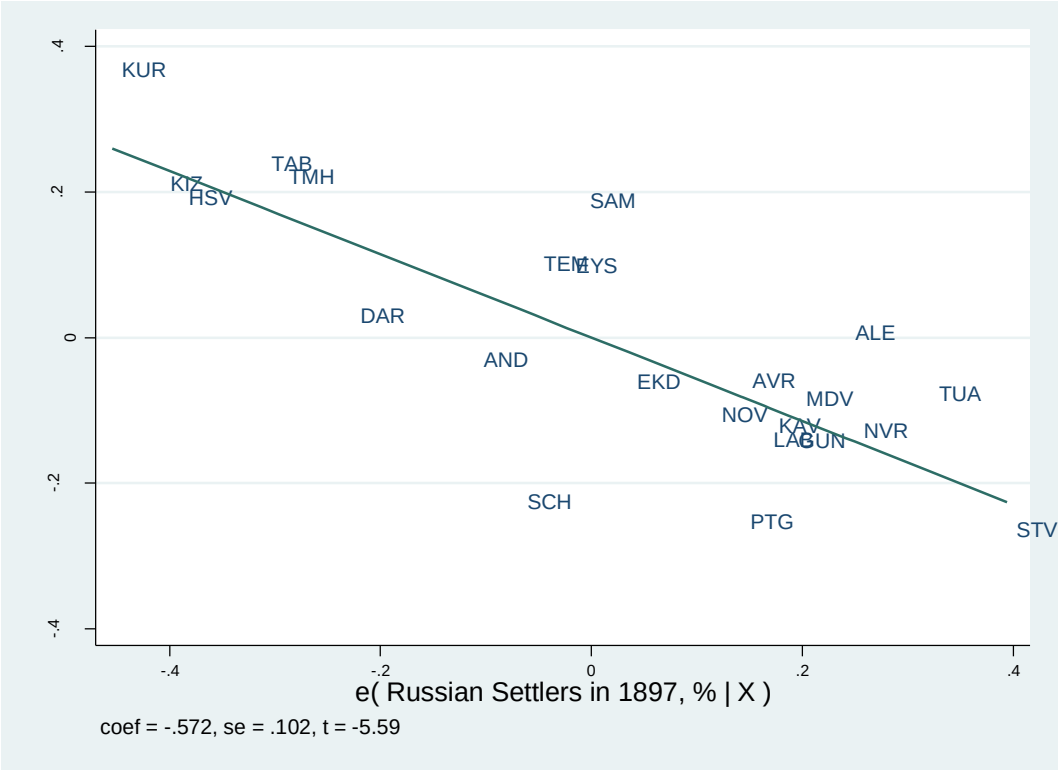
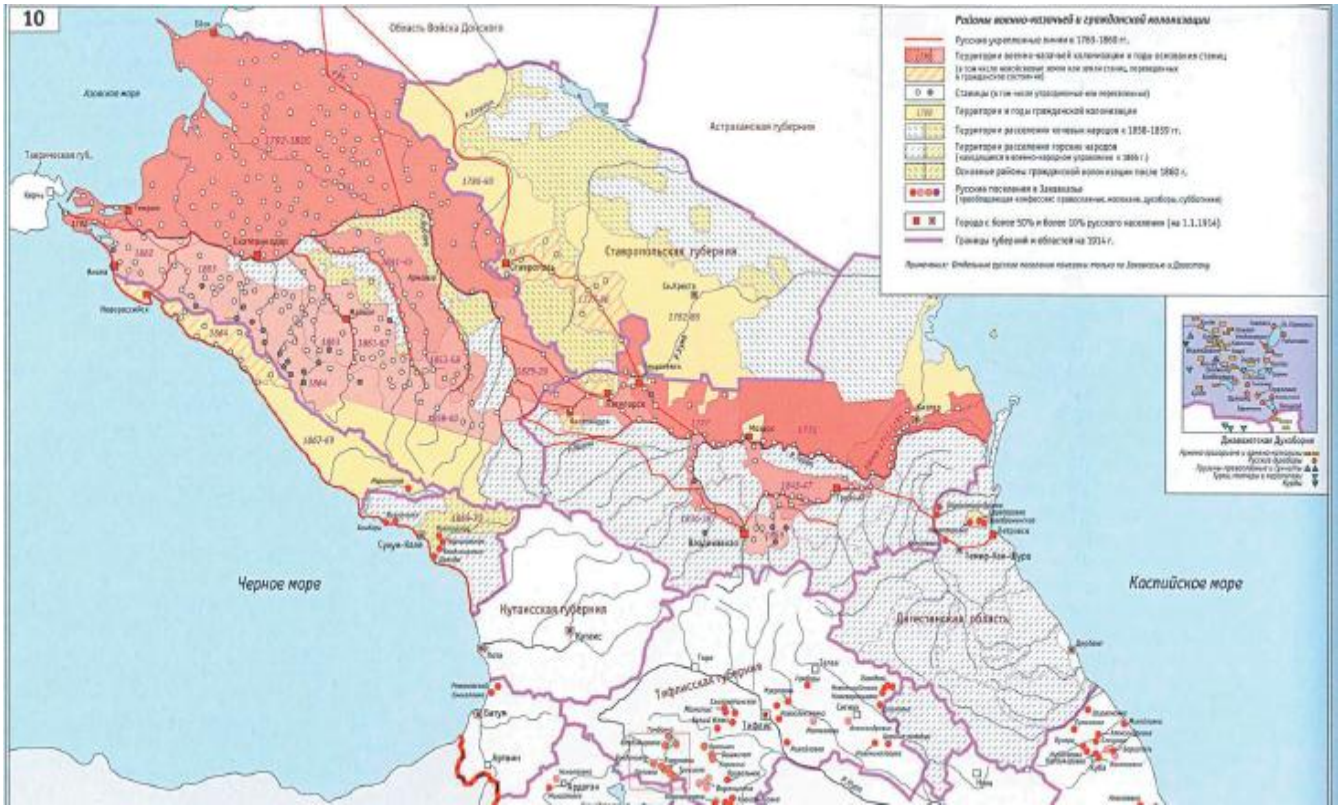


Figure 8. Russian settlement in 1897 and share of federal subsidies in local budget in 2007.



Map 1. Russian military and civil colonization of the North Caucasus.

Source: Tsutsiev (2006)

Legend:

White dots represent Russian settlements.

Red – Areas controlled by Russian military and Cossacks in the middle of the nineteenth century.

Yellow – Areas of civil colonization by Russian peasants in the late nineteenth century.

Dashed (white and yellow) – Areas of indigenous population.

Photo 1. “Monument to a Russian teacher” in Makhachkala (capital of Dagestan).