Akan Kadyrbekov[1], Dmitry Veselov[2]

Migration of Russian settlers and the legacy of interregional development in Kazakhstan[3]

Abstract

The paper explores the effect of migration of Russian settlers on the intra-regional development in Kazakhstan. We use the 1897 census dataset of the Russian Empire and modern economic data to provide links between the density of the Russian population in Kazakhstan and the current level of economic development. Exploiting exogenous geographic and geopolitical sources of variation across twenty-six districts (*uyezd*) we provide the empirical evidence of positive impact of the migration of Russians in XVIII-XIX centuries on the current level of development. The paper discusses several channels of such influence: human capital formation channel and the Soviet Union industrialization policy.

JEL Classification: N13, N33, O1, O15

Keywords: Intra-regional development, migration flows, historical development.

[1] Akan Kadyrbekov, research intern, International Laboratory for Macroeconomic Analysis, National Research University Higher School of Economics.

[2] Dmitry Veselov, senior research fellow, International Laboratory for Macroeconomic Analysis, National Research University Higher School of Economics, dveselov@hse.ru

Introduction

At the present there is a rich and still growing literature studying the causal effect of colonization on current level of economic development[5] (Acemoglu et al. 2001, Engerman Sokoloff, 1997, Nunn, Wantchelon, 2011, Natkhov, 2015). The colonial period provides an interesting natural experiment in economics because migration flows of settlers during this period are connected with geographical and climatic factors, which are exogenous by nature. The Russian Empire had a unique pattern of colonization – its expansion 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 peoples created different incentives for the metropole and, as a result, a different set of institutions and policies in the colonies.

Among territories that were conquered by the Russian Empire, Central Asia, Caucasus and Far East are the regions that significantly differ from Russia in terms of cultural heritage, informal institutions, language etc. However, unlike the other two, Central Asia has one of the most thriving economies, the most successful being Kazakhstan. There are several reasons why Kazakhstan drew our attention.

Firstly, it can be observed from the map of Kazakhstan that major cities are situated peripherally, extremely distant from each other and almost all of them were founded during the period of Russian colonialism. Moreover, the level of industrial development, concentration of educational organizations and human capital levels vary largely across the country. Can conventional wisdom explain why some regions of Kazakhstan are better developed meanwhile others have been abandoned? Or was this process predominantly influenced by the history of colonial period?

European powers are thought to have molded the way some regions of the world look nowadays. Some researchers emphasize that inflow of technologies, knowledge and foundation of educational and industrial units contributed to economic development of colonies (Gennaioli et al. 2013). Others prove that the effect of colonization on long term economic performance depends on whether extractive or inclusive institutions were established by the metropole in the colony (Acemoglu et al. 2001). The case of the Russian colonization of the Kazakh khanate remains relatively unexplored.

---

[5] A recent survey of the effect of historic events on modern economies is given in the Handbook of Economic Growth, Nunn (2014).

Current intraregional differences of economic development in Kazakhstan which can be observed in numerous examples might be the consequence of certain policies implemented by the colonial power. In this paper, we study a historical case – the colonization of the Kazakh khanate by the Russian Empire which started in the early eighteenth century that provides an opportunity for the empirical investigation of the effect of Russian settlements on local populations and the subsequent differences in interregional development of the Republic of Kazakhstan (RK), the 9th largest country in the world.

We base our argumentation on the 1897 census dataset of twenty-six districts. The territories of these districts currently belong to the Republic of Kazakhstan. In addition to this, we use data of one-hundred-ninety-nine districts of Kazakhstan in accordance with the modern administrative division of the republic. We demonstrate that Russian colonizers following its own interests created new educational institutions, which resulted in human capital accumulation among the native population. Furthermore, this effect had long-term consequences for the income levels and industry advancement in the districts (the second, we believe, is more appropriate to use as proxy of economic development, because CIS – Commonwealth of Independent Countries – use large interregional transfer programs, which distort the reliability of variables such as income per capita and average salary).

However, observed robust correlations between Russian settlement, literacy rate among Russians and long-term growth of districts raise a natural question whether settlers, and particularly literate ones, selected districts based on the literacy rate and level of initial development of indigenous people. More formally, settlement might correlate with unobserved district characteristics that have been already leading indigenous societies towards different economic outcomes. There are no available numerical resources disclosing initial characteristics of the Kazakh khanate before 1720s, however, we find the ground on the fact that Kazakh people of that time adhered to nomadic lifestyle based on non-sedentary herding (cattle breeding), except for few towns in the south. 1720s is a special decade in the history of colonization of Central Asia since the first outposts that grew to become major cities of Kazakhstan were founded that time. Therefore, Russian settlers did not select 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 colonization of the Kazakh khanate for the Russian Empire was a part of a larger primary purpose of pursuing deeper into the continent to establish and protect its southern frontier from its constant rival – the British Empire. Conquering whole Central Asia down to the Pamir mountains required colonizing the Kazakh khanate as the first step. At that time Russians and Kazakhs appeared to have common rivals –

Asian khanates, Jungaria, and China. To protect the border from them, it was essential to control the long border with no natural obstacles. Thus, the Russians' proceed into the Kazakh land followed a clear geographical pattern – northern and western regions were much more extensively settled by illiterate Russians meanwhile regions on the southern and eastern borders accounted for more literate settlers who were, probably, sent there for administrative and military purposes. As a result, the inner regions of Kazakh land did not stand out at the end of XIX century in either literacy rate among Russian settlers or the share of Russian settlers. This difference caused a substantial cross-district variation in the share of Russian settlers and literacy rate among them that were exogenous to the local population. Indeed, the only variables that predict the share of Russian settlers in the region are distance to the Russian Empire, ground access to a rival, and presence of resources – all the factors are statistically significant.

As geographical factors are exogenous by nature, the one concern is that the same factors can influence the intraregional development in Kazakhstan without the migration of Russian settlers. To state that these exogenous factors influenced human capital through Russian settlers we need to tackle the endogeneity problem. Proximity to the foreign countries provides trade opportunities that encourage the accumulation of human capital and economic development. However, in nomadic state the trade opportunity is less developed. In reasoning this issue, we can also address the non-sedentary lifestyle of indigenous people across vast majority of the Kazakh khanate's land, particularly between Irtysh, Ural, and Syrdarya rivers. Moreover, even though in early eighteenth century international relationship between the Kazakh khanate and its neighboring countries was often rival.

To tackle the problem of endogeneity we use 3sls method. Firstly, we use distance to Russia, dummy for presence of enemies or resources and their interaction to instrument number of literate Russians. Secondly, this instrumented variable along with other controls is used to instrument share of Russian settlers. Finally, instrumented share of Russian settlers is used as a factor to predict the industrial product per capita in 2000 and 2016.

To the best of our knowledge the only paper that empirically investigates the impact of colonization in the context of the 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.

Natkhov (2014) considers the effect of Russian settlers on the development of informal institutions and social advancement in the Caucasus region. The researcher executed placebo test

to show that in the North Caucasus the proximity to the sea shore determined the share of Russian population by the time of the census, meanwhile in the South Caucasus correlation between these two variables was insignificant. Natkhov proves that distance to the Black Sea negatively affects the proportion of Russian settlers in a district. The higher the share of Russians was, the lower the fertility rate was, the higher GDP per capita. It also influenced other essential determinants of Malthusian and post-Malthusian features. The novelty of our paper is that we found that the main determinant of product of industry per capita in 2000 and 2016 is share of Russian settlers instrumented by the number of literate Russians, deviation of which is thoroughly explained by the distance to Russia, closeness to potential enemies and presence of resources. In addition to this, we focus on economic indicators. Our research is the first attempt to find a causal relationship of Russian settlers and modern economic development of Kazakhstan.

The remainder of the paper is structured as follows. In the next section, we provide a short historical overview of the Kazakh khanate and the causes and manner of colonization by the Russian Empire. After that, we describe the data used to estimate the effects of colonization. Then, we describe the main econometric model and address causality issues and examine the long-term consequences of Russian settlement.

2. Historical background

2.1. General information

The Kazakh khanate was established in 1465 and expanded to the territory that lied between the Caspian Sea to the west and the Altai mountains to the East, Siberian plain to the north and middle part of the river Syrdarya to the south. The Kazakh khanate was a predominantly nomadic, monarchic, but in some occasions politicly divided country. Historically, this region was a part of the Mongol Empire the most known ruler of which, Genghis-khan, had shaped new political state based on thousands, hundreds, and tens of warriors from different ethnic groups. His decision broke apart long-lived tribal division. Hence, after years of assimilation under his and his descendants' rule the population of the Kazakh khanate became somewhat monolithic in respect to language they spoke and culture they shared. This large country had administrative division into three parts (*juz*). Their names *Kishi, Orta, Uly* can be translated into English as Younger, Middle, and Senior (Great). Surprisingly, the names of these units seemed to signify seniority or superiority, even though there was no explicit reference, for instance, to *Kishi juz* as to less influential, powerful etc. Historians have not concluded the ultimate reason of the certain names and particular land division between them yet

(Glenn, 2011). However, the most common explanation considers routes of nomadic movements and terrestrial specifications as main reasons to divide and denies any order between these units which could have resulted in names signifying seniority. There is no tribe inside the Kazakh nation that significantly differs from the rest in language, culture, appearance etc. Dialects are not prominent. All these facts allow us to implement *ceteris paribus* assumption and evaluate exogenous effects on this society's development.

Located in the center of continent the Kazakh khanate experienced overall peaceful relationship with the Russian Empire (they even had embassies), and not friendly, but quiet relationship with the Chinese Empire. By contrast, Central Asian khanates – Bukhara, Khiva, Kokand – competed for major cities and influence which frequently led to military confrontation. However, the tensest relationship was between the Kazakh khanate and the *Jungar khanate* (*Jungaria*). The latter was established on the territory of current Mongolia and desired to seize steppes controlled by Kazakhs to expand the lands for nomadic movement as Jungars adhered roughly the same lifestyle as Kazakhs.

Politically, the Kazakh khanate was torn apart between not only three *juzes* but also smaller entities within during the most intense period of Kazakh-Jungar war (1700-1720s). Only in 1720s could Kazakhs unite and succeed in resistance against foreign invaders and after several years freed all occupied territories and took back their land. However, things were different by that time - one of European major powers has already put an eye on Central Asia. In 1726 there was the *kurultai* – the summit of Kazakh tribes leaders and khans where the most important decisions are made (one of the recent kurultais was hosted by Kazakhstan after gaining independence in 2017) – where khans reconciled the possible ways to solve the Jungar problem. One way was, obviously, to unite and finish the war, but some khans doubted the chances to win on their own, so they suggested to ask the Russian Empire for help, since other neighbors were reluctant to assist Kazakhs or were hostile to them. The kurultai decided to send ambassadors to St. Petersburg to arrange terms of military alliance against Jongars, but *Abulkhair* khan, the khan of *Kishi juz* – the closest to St. Peterburg – who was responsible for sending ambassadors changed the aim of their mission from alliance to protectorate. After empress Ann signed the document to comprise *Kishi juz* in the Russian Empire, she sent also Russian *tolmachs* (diplomats who knew *Turkic* languages and culture) under Tevkelev's leadership to bribe all the other *Kishi juz* khans to ensure that they bow to empress. *Orta* and *Uly juzes* refused to become a part of the Russian Empire and insisted on alliance, but unsuccessfully. In addition to this, Russian authorities postponed military actions against Jungars until all Kazakh khans bowed to the empire. This stimulated still independent Kazakh tribes to unite and defeat the eastern

enemy, but the process of colonization has already started by that time (the first outpost has already been built).

After being pushed away to their native lands Jungars were unable to protect themselves from major attack, but Kazakhs did not continue the war after returning all their lands – it was the Chinese empire which executed genocide (Perdue, 2009) against Jungars and wiped out the whole nation. As a result, Jungar state does not exist nowadays, only Jungar descendants live in Kalmyk republic in the Russian Federation) (Khodarkovsky, 1992). The relationship was neither rival nor friendly between *Orta juz* and Chinese empire, so the former did not need Russian protection any more. Instead, the Russian Empire pursued colonization and almost a century after Anne Ionanovna's decision to comprise *Kishi juz* Russian outposts dominated all major rivers and controlled nomadic movements in *Kishi* and *Orta juzes* (Ismagulov et al. 2010). By 1860s the Russian Empire has conquered the *Uly juz* and all Central Asian khanates. The XIX century was full of uprisings against Russian rule in all juzes and khanates. They were responses of indigenous people to violations of human rights by Russian authorities, shrinking area for nomadic movements etc. The most significant uprising (1827-1847) was led by *Kenesary Qasym* who managed to unite all Kazakhs into one state and gain independence. After long war against the Russian army *Kenesary* retreated to southern part of khanate and then to *Kyrgyz* tribes, meanwhile Russian authorities promised a great amount of money to those who capture the rebel khan. *Kyrgyzs* perceived this as an opportunity, betrayed *Kenesary* and killed him (Hiro, 1994). He was the last khan of the Kazakh khanate to rule the whole state.

After the Bolshevik revolution and the Civil war Kazakhstan joined the USSR as an autonomous part of the Russian Soviet Federative Socialist Republic, namely, the Kazakh Autonomous Soviet Socialist Republic. In 1936, when the Presidium of the communist party of the Soviet Union accepted a new constitution, Kazakhstan lost its autonomy inside the RSFSR and simultaneously exited it, becoming one of fifteen republics comprised by the new USSR. This event was extremely consequential and decided that Kazakhstan would become independent after USSR's collapse.

During the World War II a lot of Russian, Ukrainian, and Belorussian factories were relocated further in the continent, including the Caucasus, Siberia, Far East and Central Asia. Since USSR was a command economy, the decisions regarding places where new factories should have been situated were entirely made by the communist party's authorities and based on certain regional characteristics and capacities, apparently, including the number of well-educated people. The transition from command economy to free market had a detrimental effect on some factories and industrial organizations many of which continued working in XXI century.

2.2. The purpose of colonization

The primary motivation for colonization of the Kazakh khanate can be clear if we 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. However, further colonization of Central Asia suggests that the Russian Empire's purpose was to establish and strengthen the border with British colonies, control trade and foreign affairs of Central Asian states with China. Central Asia faced Russia in the north, India in the south, and China in the east.

From the historical evidence it is clear, that nations in the southern part of this region initially opposed Russian rule by all means, especially, military. So, it seems that regions of Kazakhstan neighboring with those nations were primarily settled by literate Russians who were sent there with certain purposes.

The first goal of the Russian Empire was to keep Kazakh tribes ununited, because, as historians claim sedentary mostly agricultural states (such as the Russian Empire) always struggled to conquer monarchic nomadic states and to raise chances to win in case of war they always counted on outposts across wide steppes. This also applies to the policy implemented in the Kazakh lands.

By the way, note that nomadic leaders never considered their letters to the rulers of sedentary states that expressed their willingness to become a protectorate as lawful documents which bounded them politicly and economically to the receivers of the letters. So Kazakh leaders, as well as other nomadic leaders, were only looking for resources and potential trade that they could have controlled and for which they would have served as intermediaries. By contrast, the counter-party, if it was significantly stronger and technologically more developed, frequently strived for taking over power in these states in the long term both in economic and political sense. It is important to mention that on average before the mid XVIII century export of Russian industrial production to the Kazakh steppes accounted for about 30-50% of all export of industrial production of Russian manufacturers. Meanwhile, export to China and Western Europe primarily consisted of raw materials. Moreover, until 1860s the relationship between the empire and Kazakhs was managed by the Ministry of Foreign affairs (rather than Internal affairs) and trade with Central Asia was allocated to exports.

Prudently, Russian colonizers did not attempt to destroy the institute of khans' and sultans' rule because it was obvious that indigenous people were reluctant to relinquish from khanate authority to the Russian one which they were not used to. Instead, the policy of the empire implied several directions. First, Russians took advantage on weakened khans' rule and started

ceasing northern territories. Second, they made possible for Kazakhs to move to inner lands of the empire – this proved that in some instances indigenous people saw the Empire as a more reliable ruler. Third, they bribed a lot of Kazakh authorities and established local schools which appealed to indigenous people and alleviated slow adopting of the institutions of the empire by Kazakhs. Forth, after Peter I the empire was capable of building large outposts to spread military influence over steppes which made full colonization seem inevitable.

It is clear from the data that regions, close to Russia were first to be colonized and settled by Russians, whereas southern parts lacked foreign settlers by the time of census. This can be explained by the severe need of new arable lands which Russia experienced only in the second half of the XIX century – after abolition of serfdom. The reason why a lot of peasants needed to move to newly colonized lands, like Siberia and Central Asia, was the unfair proportion in which ownership of arable land was divided between landlords and peasants – to please and convince the former to accept the reform the emperor needed to allocate more territories to them. This tightened already tough lifestyle of ordinary peasants. In addition to this, only in the beginning of XX century did the minister of economy send more Russians to the furthest corners of the empire. That is the reason why by the end of XIX century only parts of Kazakhstan that were closer to Russia were populated by large number of Russians. But these facts fail to explain why so many literate Russians have already settled in the southern and eastern parts of Kazakhstan.

To examine this in detail we need to look at data. The numerical resources considering the past of Central Asia are scarce. However, there is one particularly important – the first census of the Russian Empire (1897). It provides various data on all parts of the Empire which can help us determine major factors of Russian migration and its effect on indigenous people.

3. Data and first regression results

This study explores the data of Russian Imperial census conducted in 1897, the only nationwide census conducted in the Russian Empire, and the statistical data of modern Kazakhstan. The former was obtained from 7 regional issues (one for each region (*oblast*): Torgai, Semei, Jetysu, Aqmola, Ural, Transcaspian, Syrdarya) located on the website of the Presidential library[6], and the latter was taken from the official Kazakhstani source of data[7]. Under Russian rule Kazakhstan was divided into 7 regions that were further divided into 26 districts and the census contains information considering the population, its ethnical composition (defined in terms of native language), number of males and females, literate in each represented language – Russian and non-Russian, which allows to obtain the number of indigenous people

---

[6] prlib.ru
[7] stat.gov.kz

literate in their mother tongue and Russian language[8]. In addition to this, we extracted similar data regarding ethnical composition from modern sources as well as the industrial product to use it as a proxy for level of economic development. The latter comprises products of various fields ranging from oil and electricity to hotel and restaurant services. Therefore, it is the closest to the gross regional product (GRP). Although it omits agriculture we assume that these days after massive industrialization of Kazakhstan, having only 3% of people employed in agriculture, industrial product is a good proxy for GRP.

Also, we needed to take into account that modern administrative division is different from that of colonial period. Thus, to make numbers comparable we matched all new regions to the historical districts. Almost all administrative borders remained unchanged. Where altered we regarded the location of the economic center of the region as the key determinant. Note that some modern regions were not included in either of 7 main regions of colonial Kazakhstan.

Even though the number of observations (twenty-six districts) appears small for statistical analysis, this database has several advantages. It considers the Kazakhs, who are very homogenous in terms of culture, language, attitude to foreigners and different religions regardless of the part of Kazakhstan where they lived or what lifestyle, ranging from sedentary to nomadic, they adhered to. In addition to this, we implement outlier analysis where necessary.

Ultimately, for the development of the local economy their contribution depended only on the level of their literacy.

Table 1. Descriptive statistics

Variables	Mean	Min	Max
Number of Russian settlers	20486	527	126030
Population in 1897	151278	68555	293619
Share of Russian settlers 1897	0.121	0.006	0.498
Russians' literacy rate	0.27	0.13	0.463
Indigenous literacy rate (Russian)	0.003	0.0004	0.013
Indigenous literacy rate (non-Russian)	0.029	0.007	0.09
Share of Russian settlers 2010	0.2	0.003	0.64
Log(Industrial product per capita 2000)	1.64	-0.64	4.4
Log(Industrial product per capita 2016)	4	1.68	6.8
Distance to Russia (km)	374.4	28	1175

---

[8] Many researches claim that the number of Russians might have been inflated due to political reasons (Eberhardt, 2003). On the other hand, number of indigenous people regardless of nomadic-sedentary division is precise because the capital wanted to know how many new servants are there in the empire.

Dummy (presence of resources[9])	0.19	0	1
Population 2016	655949	71917	2896754
Russian population 2016	155847	438	785581
Indigenous population 1897	129738	66590	278142
Dummy (presence of Enemy)	0.46	0	1
Number of literate Russian settlers	4681	193	37276
Area (square km)	99	5.9	249.8
Indigenous population density 1897	2262	301	19335
Population density 1897	2620	316	20785
Population density 2016	9133	506	44424
Dummy (Great Silk Way)	0.3	0	1
Latitude	48	41.3	54.87

The distance to Russia was calculated as the length of a direct line connecting the economic center of the district and the nearest point of the mainland Russian Empire. Dummy (presence of resourses) is equal to 1 if the region had a well-developed industrial plant and 0 otherwise.

Dummy (presence of an enemy) is equal to 1 if the region had a border with hostile states (the Kokand, Khiva, Bukhara khanates and the Chinese Empire) and 0 otherwise. Non-Russian language in our specification comprises a range of Turkic languages as all people who stated one of those as native language we referred to as indigenous population. Dummy for Great Silk Way is equal to 1 if the region lied along the path of the Great Silk Way which could have affected the accumulation of human capital in the area and 0 otherwise.

Thirdly, summary statistics depict large variation in almost every dimension – literacy rate, ethnic composition and geography (proximity to Russia and enemies). For instance, the proportion of Russian settlers ranges from less than 1% to almost 50%, literacy among the indigenous populations ranges from 0.7% to 10%, the literacy among Russian settlers ranges from 13% to 46% and population density ranges from 0.3 to 21 (number of people per square kilometers). As the following empirical analysis will show, this variation provides an opportunity to estimate the effects of settlement on regional development and determinants of shares of Russian settlers in regions of Kazakhstan at 1987.

3.1. Russian settlers and literacy rate of indigenous people.

From here on we use the following notation:

NumRusLit$_i$ – a number of literate Russian settlers (in thousands),

---

[9] Resources known in Kazakhstan by the time of census included oil reserves and multiple endemic types of fish in the Caspian Sea, minerals near Aqtobe delivered to factories in mainland Russia, sand in Kazakh deserts used for building purposes etc. (Kabuldinov & Kaiypbayeva, 2012)

DistRussia$_i$ – a distance to closest Russian mainland territory (in thousand km),

EnemyRes$_i$ – a dummy for the presence of enemies or resources,

DistRussia$_i$ * EnemyRes$_i$ – a intersection of two previous variables,

ShareRus1897$_i$ – the share of Russian settlers among all population,

$\widehat{\text{NumRusLit}}_i$ – fitted values of NumRusLit$_i$ from the first step,

$\widehat{\text{ShareRus1897}}_i$ – fitted values of ShareRus1897$_i$ from the second step,

LogIndust2000$_i$ – natural logarithm of industrial product per capita in 2000,

LogIndust2016$_i$ – natural logarithm of industrial product per capita in 2016.

We begin by examining the relationship between Russian settlement and the literacy of the indigenous population in 1897. The logic here is straightforward, but this stage is important to show the base on which we proceeded with our argumentation. Our baseline estimating equations are

$$\text{IndigLitRus}_i = \beta_0 + \beta_1 \text{ShareRus1897}_i + \beta_2 \text{RusLit}_i + \gamma X_i' + \varepsilon_i \ (1)$$

$$\text{IndigLit}_i = \beta_0 + \beta_1 \text{ShareRus1897}_i + \beta_2 \text{RusLit}_i + \gamma X_i' + \varepsilon_i \ (1.1)$$

where IndigLit$_i$ and IndigLitRus$_i$ are the literacy rates among the indigenous population in district I in non-Russian and Russian languages correspondingly, ShareRus1897$_i$ is the share of Russian settlers in the district in 1897[10], RusLit$_i$ is the literacy rate among Russian settlers and X_i' is a vector of control variables that, possibly, also relate to the literacy rate of the indigenous people.

Table 2 reports the estimation results for Eq. (1). The first column illustrates the regression with only one factor. It can be seen that Share of Russian settlers alone describes more than 40% of the dependent variable's variance. The second column shows that literacy rate among Russian settlers is statistically insignificant. The third column omits the share of Russians to show whether exogenous geographical factors correlates with the dependent variable. Here, being situated on the Great Silk Way is not significant. By contrast, the effect of latitude is positive and highly significant which implies that the more distant from the equator the region was the more literate indigenous people were. Nevertheless, if share of Russian settlers is included in the regression (the forth column) latitude and the GSW become insignificant.

Table 2.

Dependent variable: Literacy rate of indigenous
people (non-Russian language)

---

[10] In appendix you can find same regressions

	I	II	III	IV
Share of Russian settlers	0.0133***	0.014***		0.046**
	(0.0029)	(0.0033)		(0.022)
Russian settlers' literacy rate		0.0024		
		(0.0048)		
Great Silk Way			0.004	-0.0006
			(0.007)	(0.007)
Latitude			0.003***	0.0016
			(0.0009)	(0.001)
Constant	0.015***	0.008	-0.12***	-0.06
	(0.0052)	(0.016)	(0.043)	(0.051)
Number of observations	26	26	26	26
R-squared	0.41	0.47	0.47	0.56
Adjusted R-squared	0.39	0.42	0.43	0.5

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

One concern is native self-selection of indigenous people in regions with the presence of Russians. Aldashev (2012) shows that the migration of Russians made indigenous population turn to a sedentary lifestyle, which hade welfare reducing effect for indigenous population, especially, women in poor households. This may be one of the ways how sedentarization of Kazakhs happened and why indigenous people decided to settle near Russians. This work support the idea that the was not a self-selection of indigenous people.

In the table 3 you can find results of regressions for the equation (2). The first column reports simple regression results including only the first factor without any controls. In column (2), we account for both first factors indicated above also without any control, i.e., the share and literacy rate of the settlers. In both specifications, the estimated relationship between Russian settlement and the indigenous people's literacy is positive and statistically significant. In column (3), we control for the exogenous geographical factors. Even though, here latitude is significant, column (4) shows that when share of Russians is included geographical factors are insignificant. This will be investigated in more detail further in the 3sls. In column (5) we include dummies for oblasts Semei and Torgai. To justify dummies for two oblasts it is necessary to understand the history of that regions. Firstly, Semei region contained the capital city (Semipalatinsk) and is a

homeland of a famous teacher Abai Kunanbai (1845-1904). Torgai oblast also was a homeland of the other Kazakh teacher Ybyrai Altynsarin (1841-1889). Both opened and managed a great number of schools (second half of the XIX century).

Table 3.

Dependent variable: Literacy rate of indigenous people (Russian language)

	I	II	III	IV	V
Share of Russian settlers	0.074***	0.076***		0.012***	0.03**
	(0.018)	(0.021)		(0.004)	(0.014)
Russian settlers' literacy rate		0.006			
		(0.031)			
Latitude			0.0004***	0.00008	
			(0.0001)	(0.0001)	
Great Silk Way			0.0007	-0.0003	
			(0.0014)	(0.0013)	
Region - Semei					0.008**
					(0.003)
Region - Torgai					0.016***
					-0.0045
Constant	0.02***	0.018*	-0.018**	-0.002	0.016***
	(0.003)	(0.0099)	(0.0085)	(0.009)	(0.0026)
Number of observations	26	26	26	26	26
R-squared	0.41	0.43	0.31	0.51	0.8
Adjusted R-squared	0.39	0.4	0.25	0.45	0.76

Figure 1. The share of Russian settlers in 1897 and the literacy level of indigenous people in 1897 for each district
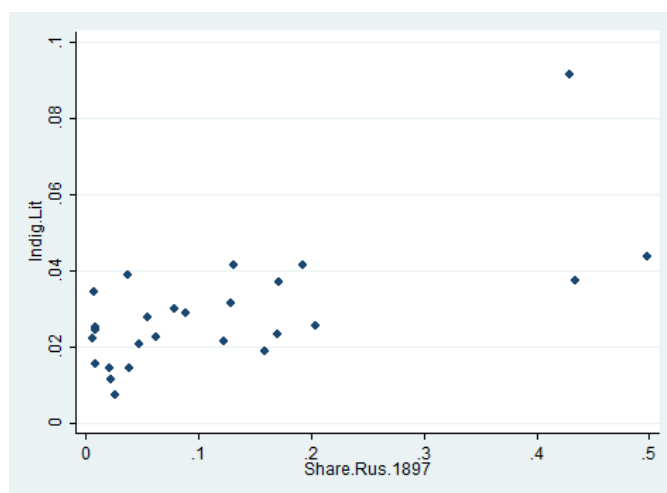
Figure 2. The share of Russian settlers in 1897 and the literacy level of Russian settlers in 1897 for each district
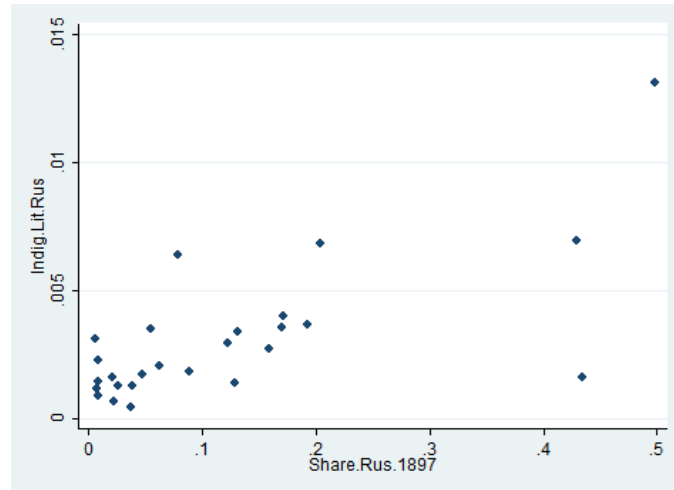
It is clear from the scatter plots above (Figure 1 and 2) that a higher share of Russian settlers among all population of the region is, in general, associated with a higher literacy rate of indigenous people in both languages. We can also see some outliers which might make the regression results vulnerable. To tackle this we implement the outlier analysis by running same regressions excluding several observations.

For regressing literacy rate of indigenous people in non-Russian languages we exclude observations with share of Russians exceeding 30% (Table 2.1). For regressing literacy rate of indigenous people in Russian languages we exclude observations with share of Russians exceeding 30% and/or with indigenous literacy rate more than 0.5% (Table 3.1). In the subsample without outliers we have the same regression results.

Table 2.1

	Dependent variable: Literacy rate of indigenous people (non-Russian language)			
	I	II	III	IV
Share of Russian settlers	0.06**	0.04		0.03**
	(0.028)	(0.03)		(0.011)
Russian settlers' literacy rate		-0.025		
		(0.023)		
Great Silk Way			-0.001	-0.002
			(0.0045)	(0.005)
Latitude			0.002***	0.0015
			(0.0006)	(0.001)
Constant	0.02***	0.029***	-0.057*	-0.05
	(0.003)	(0.008)	(0.03)	(0.031)

16

Number of observations	23	23	23	23
R-squared	0.19	0.23	0.48	0.52
Adjusted R-squared	0.15	0.15	0.43	0.44

Table   3.1

	Dependent variable: Literacy rate of indigenous people (Russian language)			
	I	II	III	IV
Share of Russian settlers	0.011***	0.016***		0.01***
	(0.0024)	(0.003)		(0.003)
Russian settlers' literacy rate		0.006***		
		(0.002)		
Latitude			0.0001	-0.000001
			(0.0007)	(0.00008)
Great Silk Way			-0.0002	-0.0004
			(0.0005)	(0.0005)
Constant	0.001***	-0.0006	-0.002	0.002
	(0.0003)	(0.0007)	(0.005)	(0.004)
Number of observations	21	21	21	21
R-squared	0.46	0.63	0.12	0.5
Adjusted R-squared	0.44	0.6	0.02	0.42

As the correlation between does not show to us the causal effect of emigration of Russians, we focus now on factors that attracted Russians to particular places and how the effect of Russian settlers persisted through centuries. The migration of Russians from the mainland metropole to the newly colonized territories seemed to happen in two stages. Firstly, administrative and military workers along with construction workers were sent by the authorities of the Empire in a directive manner to explore newly added lands. The purposes of their trips (and, eventually, a full family move) included establishing sedentary towns, constructing basic infrastructure, exploring arability of lands and administering indigenous population. Secondly, ordinary Russians migrated from their later occupations to Central Asia and Siberia, which was alleviated by the abolition of serfdom (D. Moon, 2001). Place that appealed to them were, of course, primarily, those which were already prepared for living and trade, namely, where the first wave of migrants established proper living conditions. Plenty of historical sources say in favor of this proposition (Ross, E. D., Skrine, F. 2004, Brower, D. A., 2003 among others). A lot of Russian expeditions ended up establishing small towns which then attracted both ordinary Russians and indigenous people. Clearly, by the time when the census was taken the second

wave was ongoing (Figure 3). This can be proved by the astoundingly high literacy rate among Russians who lived very far from Russia, mainly close to southern and east-southern borders of Kazakhstan. Literacy rate among Russians was in average 12% higher in those places than in the other (33% against 21%). Whereas in former the total Russian population was on average 4 times lower than in latter. This implies that by 1897 new governance of towns and outposts have been established across Kazakhstan, however, the farthest territories have still lacked illiterate Russian settlers.

Figure 3. Distance to Russia and the number of literate Russians (in thousand) for each district in 1897
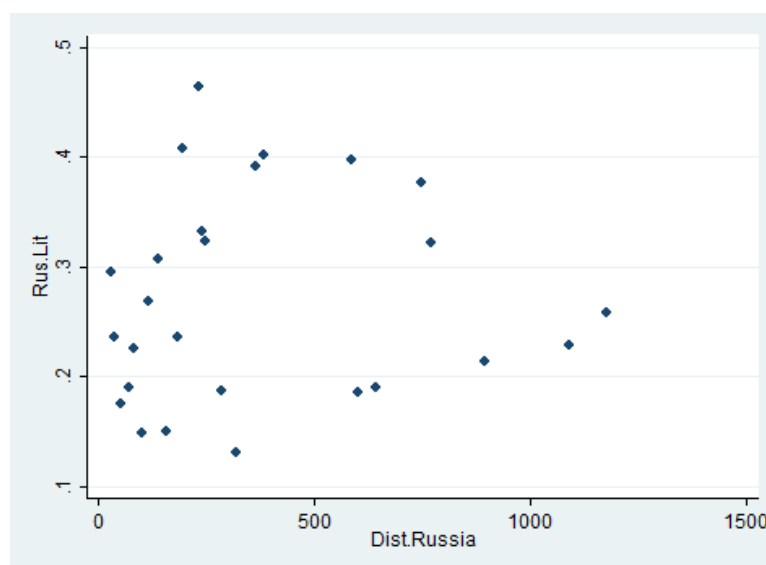
3.2. Migration of Russians and economic development of Kazakhstan. IV approach.

Historical evidence suggests that the Russian Empire had a logic of colonization closely related to the system of the Empire as a whole – many decisions were made by the leader or elites. Russian expeditions inspired by authorities or elites frequently involved highly educated people to establish the emperor's rule wherever they arrived. Moreover, the place where to settle were also rigorously chosen by these expeditors. Only literate, loyal people were considered having a chance to be a part of those expeditions and many of them also appeared in the Kazakh steppes during the first decades of colonization. This is called the first flow of migration as described previously. We realize that there is a problem of endogeneity. Only after the initial requirements for living conditions were established, could a town attract ordinary citizens and peasants. To take this key issue into account we build a 3sls model to track the causality of both flows of migration and their long-term effect on the development of Kazakhstan's regions now.

We assume that exogenous geopolitical and geographical factors, but not the initial development of indigenous people were the key factors that affected specific location of the migrants of the first flow – the literate Russians. We also provide a basis for this assumption. According to the section "Historical background", Kazakhs adhered to a non-sedentary lifestyle based upon cattle breeding, but not farming prior to the migration of Russians, with a few exceptions for the southern regions. This means that the accumulation of human capital was insignificant in those regions, in terms of the literacy rate of indigenous people and could not be regarded as a gravitating spot for migrating Russians.

The three steps of OLS are as follows:

I. $\text{NumRusLit}_i = \beta_0 + \beta_1\text{DistRussia}_i + \beta_2\text{EnemyRes}_i + \beta_3\text{DistRussia}_i * \text{EnemyRes}_i + \beta_4\text{logNumIndig}_i + \beta_5X'_i + \varepsilon_i$ (2)

II. $\text{ShareRus1897}_i = \beta_0 + \beta_1\widehat{\text{NumRusLit}}_i + \beta_2X_i' + \varepsilon_i$ (3)

III. $\text{LogIndust2000}_i = \beta_0 + \beta_1\widehat{\text{ShareRus}}1897_i + \beta_2\text{DummyRes}_i + \beta_3X'_i + \varepsilon_i$ (4)

As part of the third step, we also consider the following equation to investigate the persistence of the Russian settlers' effect, when national statistics were initially available right after the Soviet Union had collapsed and the most recent period:

$\text{LogIndust2016}_i = \beta_0 + \beta_1\widehat{\text{ShareRus}}1897_i + \beta_2\text{DummyRes}_i + \beta_3X_i' + \varepsilon_i$ (4.1)

We use the literacy rate of 1897 as an indicator of human capital and industrial product per capita in the XXI century as a proxy for the level of current development.

The first step is to evaluate regression with the equation (2). As indicated previously, the enemies were the three major khanates immediately to the south of Kazakhstan, and China was considered a potential enemy. In addition to this, various natural resources were found in Kazakhstan, some of which required considerable effort of the Empire to exploit.

Table 4.

	I	II	III	IV	V	VI	VII
			Dependent variable: Number of literate Russian settlers in 1897				
Distance to Russia	-6766*	-6729	-21510***	-26520***	-28683***	-26447***	
	(4329)	(4425)	(9882)	(9374)	(9638)	(9940)	
Distance to river		1596					
		(9650)					
Distance to Russia*Dummy (Close to Enemy)			21304**				
			(9917)				
Dummy(Close to Enemy)			-9143**				
			(4263)				

19

	I	II	III	IV	V	VI	VII
Distance to Russia*Dummy (Enemy and/or Resources)				26247***	23425***	25412*	
				(8994)	(9601)	(10996)	
Dummy (Enemy and/or Resources)				-12280***	-10960***	-12181***	
				(4044)	(3842)	(4117)	
Log (Number of indigenous population)					6116*		
					(3494)		
Indigenous population density						0.2	0.036
						(0.38)	(0.42)
Constant	7214***	7007***	11693***	14869***	13656	14577***	4597***
	(0.039)	(2427)	(2975)'	(3161)	(19639)	(3269)	(1668)
Number of observations	26	26	26	26	26	26	26
R-squared	0.27	0.28	0.34	0.37	0.43	0.38	0
Adjusted R-squared	0.25	0.25	0.27	0.29	0.32	0.26	0

We can see a result that, because the intersection of $DistRussia_i$ and $EnemyRes_i$ is included, the negative influence of distance to Russia is partially offset by the presence of enemies and resources. Since resources were not used by indigenous people prior to colonization and being close to enemy did not affect positively the human capital of that region, clearly, literate Russian settlers were gravitating to certain places prompted by exogenous geopolitical and geographical features of these territories. The indigenous population density as an indicator which can serve as a proxy for initial development is insignificant. This also support the idea of Russian migration's exogeneity.

Thus, we obtain the following variable $\widehat{NumRusLit}_1$ taken from the regression V in the table 4.

In the next step we use the following equation:

$$ShareRus1897_i = \beta_0 + \beta_1\widehat{NumRusLit}_1 + \beta_2 X_i' + \varepsilon_i \ (3)$$

Table 5.

	I	II	III	IV	V	VI
	Dependent variable: Share of Russian settlers in 1897					
Distance to Russia (thousands km)	-0.16*	-0.15*	-0.38**	-0.012***		-0.7*
	(0.079)	(0.079)	(0.17)	(0.019)		(0.36)
Distance to river (thousands km)		0.17				
		(0.17)				

Distance to Russia*Dummy(Close to Enemy)		0.37**				0.42*
		(0.19)				(0.2)
Dummy(Close to Enemy)		-0.2	0.00002			0.13
		(0.079)	(0.00008)			(0.12)
Log (Number of indigenous population)					-0.117**	-0.29**
					(0.0518)	(0.111)
Fitted values of Number of literate Russian settlers					0.02***	0.041***
					(0.004)	(0.01)
Dummy(Resources)						
Constant	0.17***	0.15***	0.0006	0.0015	0.59**	1.1***
	(0.039)	(0.045)	(0.0017)	(0.0005)	(0.27)	(0.4)
Number of observations	26	26	26	26	26	26
R-squared	0.41	0.47	0.47	0.46	0.47	0.58
Adjusted R-squared	0.39	0.42	0.4	0.42	0.41	0.44

The last column of the table shows the most interesting version of the regression with control variables. It indicates the negative influence of distance to Russia which was partly offset by the presence of an enemy. It also shows a significant positive correlation with the fitted values of the previous regression, namely, the number of literate Russian settlers cleaned from endogeneity.

From this regression we take the fitted values of the share of Russian settlers in 1897.

The next step involves the two following regression equations:

$$\text{LogIndust2000}_i = \beta_0 + \beta_1 \widehat{\text{ShareRus}1897}_1 + \beta_2 \text{DummyRes}_i + \beta_3 X_i' + \varepsilon_i \quad (4)$$

$$\text{LogIndust2016}_i = \beta_0 + \beta_1 \widehat{\text{ShareRus}1897}_1 + \beta_2 \text{DummyRes}_i + \beta_3 X_i' + \varepsilon_i \quad (4.1)$$

Table 6.

	Dependent variable: log(Industrial production per capita 2000)				
	I	II	III	IV	V
Share of Russian settlers in 1897	1.89		2.1*		
	(1.69)		(1.22)		
Share of Russian population in 2010		2.51*		1.9*	
		(1.28)		(1.1)	
Dummy(Resources)			2.03***	1.85***	2.04***
			(0.43)	(0.44)	(0.42)

	I	II	III	IV	V
Fitted values of Share of Russian settlers in 1897					3.01***
					(1.4)
Constant	1.43***	1.15***	1.01***	0.97***	0.9***
	(0.3)'	(0.34)	(0.24)	(0.26)	(0.26)
Number of observations	26	26	26	26	26
R-squared	0.05	0.14	0.52	0.51	0.54
Adjusted R-squared	0.01	0.1	0.48	0.47	0.5

Table 7.

	Dependent variable: log(Industrial production per capita 2016)				
	I	II	III	IV	V
Share of Russian settlers in 1897	-0.19		0.004		
	(1.86)		(0.53)		
Share of Russian population in 2010		0.69		0.24	
		(1.47)		(1.25)	
Dummy (Resources)			1.84***	1.86***	1.91***
			(0.53)	(0.55)	(0.5)
Fitted values of Share of Russian settlers in 1897					3.8**
					(2.7)
Constant	4.05***	3.89***	3.68***	3.72***	3.2***
	(0.3)'	(0.39)	(0.3)	(0.32)	(0.34)
Number of observations	26	26	26	26	26
R-squared	0.004	0.009	0.34	0.34	0.54
Adjusted R-squared	0	0	0.28	0.28	0.5

According to the results for these regressions, the positive influence of the instrumented share of Russian settlers on industrial product per capita in both 2000 and 2016 is statistically significant, meanwhile the influence of non-instrumented share of Russian settlers is insignificant. In addition to this, we needed to consider resources that represent the major part of Kazakhstan's industry. Clearly, the results of 3sls assert that, if cleaned from the endogeneity, the share of Russian settlers in 1897, positively affects the level of economic development in Kazakhstani regions nowadays.

To prove that the human capital mechanism is one of the main driving force of our results we need also to find out if the current industrial product correlates with the human capital during colony era. To do that we can run a two-step IV regression.

$$1^{st} step: \widehat{IndigLit}_i = 0.15^{***} + 0.13^{***}ShareRus1896_i$$

$$2^{nd} step: LogIndust2000_i = 0.5 + 26^{**}IndigLit_i + 2^{(***)}DummyRes_i$$

From that, you can see that economic development of regions positively correlate with the literacy rates once achieved. For this reason, above 3 steps least squares approach is sensible to find how exogenous factors led to specific economic development.

4. Conclusion

Our research explores the link between the two-stage Russian migration into the Kazakh steppes and the modern interregional divergence in levels of development that persist nowadays. Clearly, the share of Russian settlers among the population of the region was positively correlated with the literacy rate of the indigenous people in 1897. The instrumented share of Russian settlers is highly significant and positively correlates with the industrial product per capita in both 2000 and 2016, i.e. right after the Soviet Union collapse and currently, which justifies the persistence of the Russian settlers' effect on the indigenous people. Add to that the logic of the Soviet administration on placement of particular industrial plants and building of infrastructure and you will get a cohesive picture of Kazakhstan's development pattern.

Further research should consider other indications of the level of development of the past and of the present. Instead of literacy rate and industrial product per capita, one can take the fertility rate, or an average size of a family, or the proportion of workers employed in certain fields and occupations.

The results of these tests suggest that our main hypothesis is plausible – the distance to the main-land Russian Empire, the distance to the southern and eastern enemies influenced the literacy of the indigenous population and further, the level of human capital accumulation and industrial development through Russian settlements and not through omitted variables. Like OLS coefficients, the IV coefficients are significant. The settlements effect had long-term consequences for local development. Districts that had higher shares of the Russian settlers in the late nineteenth century, today, have larger industrial production per capita, higher educational attainment among the indigenous populations. Overall, the evidence supports the impact of Russian settlement in the Kazakh khanate on long-term development of regions in Kazakhstan during both the late nineteenth century and the contemporary era.

Finally, we explore several potential mechanisms of influence – school building, social structure, and infrastructure. During the Soviet era, a lot of decisions were made in favor of

establishing specialized schools and universities in regions with a higher Russian population share due to the greater capacity of the Russian/Soviet state to implement new rules and regulations in the settled districts. This, certainly, affected the process of accumulation of human capital which persisted after the USSR's collapse and molded a new middle class.

For this reason, Russian colonization of Kazakh lands was an event that changed not only the ethnic composition of the population but also the occupational diversity and industrial structures in the region, resulting in divergent development paths for more settled and less settled districts. Meanwhile, railroad construction could not have influenced interregional differences as well as Russian settlements distribution, by 1897, because the first rail road transportation started only in 1906.

Acknowledgments

5. References

Acemoglu, Daron, Johnson, Simon, Robinson, James, 2001. The colonial origins of comparative development: an empirical investigation. The American Economic Review 91, 1369–1401.

Aldashev, Gani, and Catherine Guirkinger. "Deadly anchor: Gender bias under Russian colonization of Kazakhstan." *Explorations in Economic History* 49.4 (2012): 399-422.

Brower, D. A. Turkestan and the Fate of the Russian Empire. London: Routlege, 2003

Eberhardt, Piotr. Ethnic groups and population changes in twentieth-century Central-Eastern Europe: history, data, analysis. Translated by Jan Owsinski. M.E. Sharpe, 2003

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.

Gennaioli, Nicola, La Porta, Rafael, Lopez de Silanes, Florencio, Shleifer, Andrei, 2013. Human capital and regional development. The Quarterly Journal of Economics, 105–164.

Glenn, Curtis, 2011. In the Soviet Union. Kazakstan: A Country Study. GPO for the Library of Congress. Retrieved 19 February 2011.

Hiro, Dilip, Between Marx and Muhammad: The Changing Face of Central Asia, Harper Collins, London, 1994,

Ismagulov, Orazak & Ismagulova, Ainagul, 2010, Physicaal anthropology of Kazakh people and their Genesis. Ch., Valikhanov Institute of History and Ethnology. Almaty, Kazakhstan.

Kabuldinov, Z.E. & Kaiypbayeva A.T., History of Kazakhstan (XVIII century – 1914). 2nd edition, 2018, - 328 p.

Khodarkovsky, Michael. Where Two Worlds Met: The Russian State and the Kalmyk Nomads 1600–1771, Cornell University Press, 1992.

Moon, David, "The Abolition of Serfdom in Russia". Harlow: Pearson Education Limited, 2001. Page 37

Natkhov, T. (2015). Colonization and development: The long-term effect of Russian settlement in the North Caucasus, 1890s–2000s. Journal of Comparative Economics, 43(1), 76-97.

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.

Perdue, Peter C (2009). China Marches West: The Qing Conquest of Central Eurasia (reprint ed.). Harvard University Press.

Ross, E. D., Skrine, F. H. B. The Heart of Asia: A History of Russian Turkestan and the Central Asian Khanates from the Earliest Times. London: Methuen & co., 1899; London — New York: Routlege, 2004.

Shubina, Elena, Aldashev, Gani, Henry, Sabine, 2014. From Saddlers to Harrows: Agricultural Technology Adoption during Russian Colonization in Kazakhstan. University of Namur, Mimeo.

Authors:

1. Akan Kadyrbekov, research intern, International Laboratory for Macroeconomic Analysis, National Research University Higher School of Economics, akankadyrbekov@gmail.com

2. Dmitry Veselov, senior research fellow, International Laboratory for Macroeconomic Analysis, National Research University Higher School of Economics, dveselov@hse.ru