Nomadic Pastoralism, Colonization and Conflict in Central Asia

Pavel Bacherikov[†]
UC Berkeley Haas

November 12, 2024

Job Market Paper
[Click here for the latest version]

Abstract: This paper studies how a sharp exogenous increase in land pressure resulting from massive land expropriations and in-migrations of peasant-settlers organized by the Russian colonial authorities in the late 19th—early 20th century affected social structures and economic activities of indigenous nomadic pastoralists in Central Asia. I assemble a novel household-level dataset constructed from two waves of Russian colonial censuses of nomadic population in 1896-1901 and 1908-1913 combined with hand-collected data from archival plot-level annual land expropriation reports that, together with landuse-based expropriation rule, allow me to use fuzzy regression discontinuity design to show that those nomadic households that experienced expropriations between ca. 1897 and ca. 1908 were more likely to partially sedentarize and intensify the use of the most fertile lands they were left with. Within extended households and beyond, an increase in land pressure facilitated the development of more individualized ownership and use rights for land, as well as gave rise to contractual labor market and rental market for land. Such a shift from pastoralism to semi-sedentary mode of production rapidly reduced the importance of top-level clan and tribe institutions traditionally regulating the use of common pastures. Instead, lower level sub-clan self-identification became more salient, households started to invest more in agricultural tools and construction of permanent buildings.

[†]Email: pbacherikov@berkeley.edu. I am especially indebted to Guo Xu, Ernesto Dal Bo, and Barry Eichengreen for their advice and support. I am grateful to Ran Abramitzky, Dias Akhmetbekov, Sebastian Arechaga, Silvia Barbareschi, Noah Bender, Volha Charnysh, Eva Davoine, Brad J. DeLong, Silvia Farina, Morgan Foy, Saumitra Jha, Gabriel Granato, Samuel Kapon, Otto Kienitz, Igor Kolesnikov, Sara Lowes, Lukas Leucht, Andrei Markevich, Tamri Matiashvili, Eoin F. McGuirk, Christopher M. Meissner, Nathan Nunn, Miguel Ortiz, J. Landin Smith, Steven Tadelis, Francesco Trebbi, and Jonathan Weigel for insightful comments and suggestions. Seminar audiences at UC Berkeley, Stanford, and Columbia, and attendees at the Berkeley Political Economy conference, Bay Area Economic History Workshop, and All-California Economic History Job Market Conference provided valuable feedback. The research for this project has been generously supported by the Institute for Business Innovation at UC Berkeley. Sonya Chervinski and Dmitry Padyshev provided excellent research assistance.

1 Introduction

Economic historians, development economists and other social scientists have long been debating the role traditional institutions—in particular, extended families, kinship networks, clans, etc.—play in shaping individual incentives and determining aggregate economic outcomes and development paths (Granovetter 2005). Kinship networks are found to be fundamental in reducing transaction costs and facilitating trade, especially in environments where formal institutions are weak or absent (Greif 1989, 1993, 1994). Together with extended families, they create social capital that enhances trustworthiness and facilitates cooperative behavior (Coleman 1988), promote engagement in resource-pooling, risk-sharing and communal consumption-smoothing arrangements (Platteau & Abraham 1987, Rosenzweig & Stark 1989, Townsend 1994, Ensminger 1997, Fafchamps & Lund 2003), improve access to credit by reducing information asymmetries and enforcing repayment (Udry 1994, LaFerrara 2003, Gupta 2014), and reduce cost of labor search and improve labor market outcomes and productivity (Wegge 1998, Munshi 2003, Foster & Rosenzweig 2010, Leunig et al. 2011). However, overreliance on these informal institutions might lead to underinvestment in human capital and, ultimately, hinder assortative matching on the labor markets (Munshi & Rosenzweig 2006, Greif & Tabellini 2017, De la Croix et al. 2018).

What is oftentimes missing in these discussions is the question of whether these institutions are able to transform in response to external treatment, and to what extent these hypothetical changes, in turn, would facilitate or hinder adaptation. In fact, there is a rich literature documenting high degree of persistence of these social structures (Acemoglu et al. 2001, Leonardi et al. 2001, Nunn n.d., Dell 2010, Guiso et al. 2016). In line with this premise, development economists tend to postulate that culture and traditional institutions such as marital arrangements, inheritance customs, co-residence practices, etc. change slowly and that, therefore, one could assume them fixed when examining the impact of various policies or shifts in endowments. Several recent studies have challenged this view in the context of household composition change in response to a technological shock (Foster & Rosenzweig 2002), land policies (Bardhan et al. 2014), and rising land pressure (Guirkinger & Platteau 2015, Guirkinger & Aldashev 2016, Aldashev & Guirkinger 2017). Nonetheless, the simul-

taneous evolution of culture and traditional institutions to accommodate for or in response to socio-economic changes remains understudied in economics, mainly because of the lack of data corresponding to events of large-scale economic change.

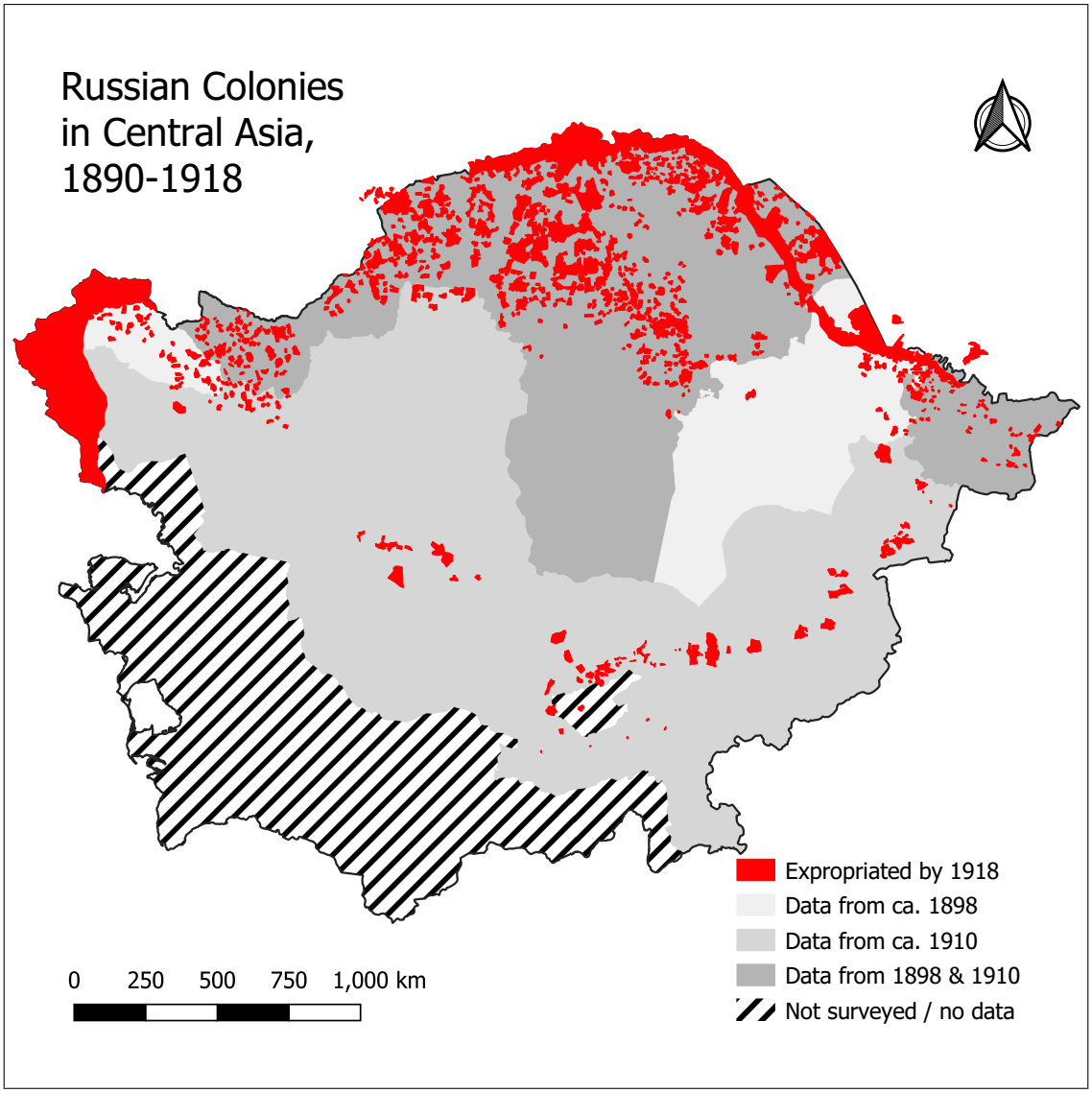
In this paper I study one of such episodes for which I've been able to collect a unique panel data containing both the measures of evolving social structures and behavioral outcomes: traditional kinship networks of nomadic pastoralists in Central Asia amid massive land expropriations and in-migrations of millions of peasant-settlers organized by the Russian colonial authorities in the late 19th and the early 20th centuries. In particular, using the data from two waves of Russian colonial censuses of nomadic population in the region, I am analyzing how changes in traditional rules governing the allocation of resources – land and labor – within Kazakh, Kyrgyz and Karakalpak extended families in response to an exogenous increase in land pressure have facilitated an evolution from a purely nomadic pastoralist regime to a mixed mode of production with significant share of population now being involved in sedentary agriculture.

More specifically, I combine newly digitized data from two partially geographically overlapping waves of colonial censuses commissioned by the Russian imperial Resettlement Administration in 1897-1901 and 1907-1913 which managed to cover approximately 3.5 million people with hand-collected archival data from dozens of unpublished annual county-level[1] plots expropriation reports produced by local resettlement officers (which give me extremely granular measure of the treatment variable) to estimate the effect of an exogenous increase in land pressure on evolution of nomadic social structures and economic activities. Measuring the effect of massive land-based policy interventions, especially in a colonial context, is usually hard due to oftentimes arbitrary or endogenous nature of assignment into treatment. In this setting, land expropriations were following a transparent predetermined rule which allows for causal estimates of the effect of interest. Based on the first wave of census, the Resettlement Administration calculated that a typical nomadic nuclear family of 6 people[2] needs a livestock equivalent of 24 adult horses to survive through a normal winter.

[1]Throughout the text, I will be using the following convention regarding translation of administrative units of the Russian Empire: *zemel'naya obshchina* – commune, *volost* – sub-county, *uyezd* – county, *oblast* – province, *general-gubernatorstvo* – general-governorship.

[2]An actual per-family number of people co-habitating together through the winter was 5.93 based on the 1897-1901 extended household-level data from three provinces.

Figure 1: The Extent of Russian Colonization of Central Asia by 1918



Note: The map depicts all land plots (*in red*) expropriated by the Russian colonial administration—sporadically by local authorities between 1890 and 1896 or earlier, and by the survey parties sent by the Resettlement Administration from 1896 onwards—in Central Asia. Portions of the Ural Cossack Host's communal lands that have been gradually confiscated in the course of the 16th-late 19th centuries is on the far West of the map. Some of the lands within the 10-*versta* continuous line in the North and North East have been first expropriated in 1639 (top line stretching from Kurgan *uyezd* to Omsk city), then again in 1760-1763 (downward *Irtysh* line from Omsk to Bukhtarma and Ust-Kamenogorsk). In 1839, the lands within the line were all granted to Siberian Cossack Host, only to be used later for the construction of the portion of the Trans-Siberian railway.

Each surveyed county was then partitioned into up to 28 survey districts based on similarity of land quality, proximity of water sources, etc. Using these agro-climatic inputs, the Resettlement Administration then measured the quality-weighted amount of land required to maintain one adult horse through a typical winter under pure pastoralist regime in this district—in other words, a district-specific norm of land per one horse. Diving the quality-weighted amount the land an extended household in this district has access to[3] by this norm, and then dividing again by the number of nuclear families in that extended household, yields the number of adult horses an average nuclear family in that extended household would be able to support through a normal winter under a pure pastoralist regime. Extended households found to be able to support more than 24 horses per nuclear family were flagged for land expropriations.

However, the timing of expropriations was often subject to local bureaucratic bottlenecks such as availability of land engineers, approval by county head and local resettlement officials, etc. On top of that, expropriation rules designed by the Resettlement Administration explicitly prescribed creation of topologically continuous plots so to prevent stripholding, which sometimes prompted expropriation officers working on the ground to confiscate the lands from households that formally were not eligible for expropriations. Therefore, I don't have perfect compliance on either side of the 24-horses threshold. Nevertheless, I show that this rule causes extended households just above the landuse-threshold to be 20 percentage points more likely to be subjected to land expropriation, allowing me to estimate the causal impact of an increase in land pressure using fuzzy regression discontinuity design. On average, I observe outcomes three to four years after the initial expropriation, which means I am estimating the short-run impact of this policy.

In contrast to many evidence regarding the slow-moving nature of traditional network-based social structures documented in the literature (North 1990, Roland 2004), I demonstrate rapid evolution of nomads' self-identification within traditional clan-based social structures. Using rich geneological data collected during both waves of censuses, I show that households just above the expropriation threshold report going down one whole generation

[3]As I discuss later in subsection 2.2, all extended households have restricted-access winter stop lands and communally-used pasture lands.

along their geneological trees when talking about their self-identification, meaning that they now prefer to identify with much tighter group of people, which is a reflection of their rapid transition from transhumant pastoralist mode of production—under which they have to coordinate and cooperate with their distant relatives at the clan level when they oscillate between seasonal pastures—to semi-sedentary intensive agriculture and husbandry where production is taking place at lower—extended household and nuclear family—levels. I also document an almost 2 percentage points increase in the command of the Russian language among adult members of expropriated extended households, which I interpret as evidence of increased cooperation between settlers and nomads as a result of forced coexistence.

I do also find that expropriations led to a rapid emergence of contractual market for labor and rental market for land among the nomads. As expropriated households were transitioning to more intensive use of land remained in their possession, they have increased the amount of land they would rent from their neighbors, both for growing crops and making hay. Households just above the expropriation threshold were also more likely to both hire farmhands and send members of their own household to work elsewhere. Quantitatively, I find that households just above the 24-horses threshold increased their exposure to the two-sided rental market of land by 2 more acres (0.73 *desyatinas*) per nuclear family, and their number of wage workers—either hired in or sent away—increased by a whopping 4.4 workers more per nuclear household vis-a-vis non-expropriated households.

I demonstrate that transition to intensive agriculture induces affected households to invest more in agricultural tools. As more members of such households shift away from seasonal migrations and instead spend more time working the land and preparing fodder for their livestock, families start investing more in construction of permanent buildings. As such, households just above the expropriation threshold have 2.6 more agricultural tools per nuclear family in their possession, which includes all sorts of wagons, carts, heavy plows, manual or horse-drawn seeders, etc. Such extended households also end up having one more permanent building—storage unit, open-air kitchen, animal stall, housing unit, etc.— per nuclear family on their winter land, which is a reflection of both more individualized use of winter stop lands and a declining importance of transhumant migrations.

Transition to sedentary agriculture is expectedly associated with an increase in the area

allocated for cropmaking: expropriated household near landuse-threshold have 3.4 acres more devoted to various crops. However, since the lands suitable for haymaking do not get expropriated at the same rate as the most fertile winter stop lands, I do find an increase in the amount of land devoted to individual (that is, at the nuclear family level) haymaking, suggesting that hay is continued to be produced collectively within extended households. Not surprisingly, I also find that reallocation of factors of production towards agriculture results in the reduction of the size of herds: households just above the 24-horses threshold cut their livestock numbers by 1.45 livestock units (expressed in adult horses) per nuclear family. Note, however, that this reduction might also be due to the fact that partial sedentarization is associated with a reduction in risk of running insufficient fodder for the livestock, so families no longer have to maintain larger herds to account for higher livestock mortality. It might also, at least partially, reflect a shift in tastes towards agricultural products. Given the data that I have at present, I cannot disentangle those channels.

Finally, to flexibly estimate the effect of land expropriations on cropland and agricultural productivity, I collect data on population and cropmaking at a subcounty (*volost*) level from archival records (CSA of the Republic of Kazakhstan) and at a county level from published gubernatorial reports from 1880–1917. For each subcounty, I record the first year it experienced land expropriation and then construct a binary treatment variable equal to 1 in the year it was first expropriated and all years thereafter. Then, I use a simple event study design to estimate the effect of land expropriations on the extensive margin of land use for cropmaking and on area-weighted crop productivity estimated over five major crops. I find that an increase in land pressure prompts immediate increase in the area devoted to cropmaking, as well as an immediate decline (though the point estimates are much noisier) in crop productivity. While the first result is in line with the household-level evidence (at least in terms the direction of the effect), the second one is at odds with much of what colonial officers tasked with studying the consequences of land reallocation have reported (see Smirnov 1899, 1900, Special Commission on the Needs of Agriculture 1903). I argue that an increase in productivity reported in the literature is hard to reconcile with what was actually happening on the ground: amid land expropriations, affected nomadic families were inadvertently pushed into using more marginal lands which are, by definition, more

suitable for less productive crops. Given such drastic change in the composition of lands available to affected households, one would need to assume a crop production function with unrealistically large positive (and preferably non-decreasing) marginal products of land and labor to make an increase in productivity possible. I therefore conclude that my estimates of the change in agricultural productivity, while noisy, are at least correct in terms of the sign.

The rest of the paper is organized as follows. Section 2 provides some background information about the functioning of nomadic economy prior to colonization. It then discusses intellectual origins of the Resettlement Administration tasked with colonization efforts in Siberia and Central Asia, and then provides important details about how settler-colonization of Central Asia was done in practice. Section 3 discussed the sources used to construct all the data used in this paper. Section 4 describes the empirical strategies. Section 5 discusses the results. Finally, Section 6 concludes.

2 Historical Background

2.1 Nomadic Economy before Colonization

Prior to the massive expropriation of land and in-migration of peasant-settlers from the European part of the Russian Empire in the late 19th century, the economy of Kazakhs, Kyrgyz and Karakalpaks was primarily based on nomadic pastoralism, a system deeply intertwined with the social organization of indigenous clans and shaped by the specific agro-climatic conditions of the steppes. This intricate relationship between economic activity, social structures, and environment existed for centuries, with nomadic pastoralism becoming the dominant production system in the north of Kazakhstan around 1500-1000 BCE (Akishev 1972, Khazanov 1975, Markov 1976, Masanov 2011). Most of Kazakhstan's terrain is characterized by a semi-arid climate and vast steppes with extreme temperature fluctuations between scorching summers and bitterly cold winters. This climatic variability, coupled with scarce and unevenly distributed rainfall, made settled agriculture a precarious endeavor in many parts of the region.

This ecological context gave rise to the practice of transhumance, which involved long-distance seasonal movements of people and their livestock between distinct summer and winter pastures, with shorter stays on spring and autumn pastures. The length of these routs could stretch from 30-50 kilometers to 1000 kilometers in the north of Kazakhstan, with an average exceeding 200 kilometers (Matskevich 1929). Summer pastures offered abundant grazing opportunities during the warmer months; however, these areas often became uninhabitable during winters due to temperatures normally falling below –35 degrees Celsius during the night, which, coupled with strong ghastly winds, often resulted in large areas being covered with thick layer of ice by the morning, making gazing impossible. Meanwhile, winter pastures, endogenously chosen to be situated near rivers, lakes, ground waters, or hills that gave relative protection from harsh weather, provided access to limited but essential fodder during winter months. The existence of such lands was vital for pastoralist lifestyle as it allowed the herds to wait out extreme weather events affecting more marginal pastures that effectively rendered them unusable for extensive grazing.

Reliance on pre-existing geographic allocation of natural endowments (land, water, etc.)

made nomadic pastoralist economy inherently susceptible to environmental shocks of any kind, particularly to certain extreme weather events, known in Kazakh as "*jut*," characterized by prolonged freezing days and formation of thick ice cover over large areas of unprotected pastures. Particularly bad years with multiple *jut* events could have decimated livestock populations, leading to a sharp decrease in quality of life and, potentially, famine. To mitigate these risks, Kazakhs developed sophisticated strategies, including herd diversification (i.e., combining livestock that is resistant to different climate into a single herd), insuring against bad weather by keeping the high-quality winter stop plots fallow to make sure there is enough naturally-produced biomass beneath the snow, and developing reciprocal arrangements within their social structures, as well as adopting more intensive forms of use of lands at their disposal, such as winter haymaking with its subsequent storage.

2.2 Nomadic Family and Clan Institutions

Indigenous social structures were pivotal for the functioning and resilience of the Kazakh nomadic pastoralist economy. Clans acted as collective action vehicles built on top of extensive kinship networks based on patrilineal descent[4], providing a framework for social organization, economic cooperation, and mutual support.

The organization of the clan system was characterized by a multi-layered hierarchical structure. The smallest unit at the very bottom was a nuclear family (*tulin*), which typically consisted of a married couple with their kids, workers, and poorer relatives, who all worked together and privately owned their livestock. Next there were extended households (*aul-qystau*), comprising multiple related nuclear families that had closed-access common ownership of lands on winter stops (hence the second part of the name, *qystau*) and coordinated their movements during seasonal migrations. Extended households were led by a respected elder, the *aqsaqal*, who played a key role in decision-making, conflict resolution, and upholding customary law. Several dozens (sometimes – more) extended households were grouped into a clan (called *ata-balasy*), all members of which shared the same commonly recognized ancestor. Finally, at the tope level there were tribes (called *ru*), which

[4]A strict exogamy rule prohibited marriages within the same clan. Clan identity was typically transmitted from father to son, while married women were incorporated into their husbands' clan.

were themselves partitioned between three confederations or hordes: Junior, Middle, and Senior. These top levels of ancestral structures didn't act like governing bodies, but as collective action vehicles regulating inter-clan conflicts, managing diplomatic relationships, and mobilizing for war with neighboring peoples and countries.

The clan system was instrumental in regulating access to and managing scarce resources, particularly land and water. Winter pastures were typically treated as common property within the extended household[5], with livestock from different families grazed collectively. Summer pastures, on the other hand, were often exploited jointly by several extended families belonging to the same clan. Clan elders played a crucial role in coordinating the timing of transhumance to summer pastures, ensuring organized movement, defense against potential raids, and efficient allocation of grazing lands (Tolybekov 1971).

The clan system also served as a vital social safety net, providing a mechanism for reciprocal assistance and risk sharing in times of need. Families facing hardship, such as loss of livestock due to *jut* or other calamities, could rely on support from fellow clan members, often in the form of livestock gifts. This system of mutual aid was facilitated by the geographic dispersion of extended families within a clan, mitigating the impact of localized shocks.

The clan system's influence extended to the adoption of new technologies, particularly in the context of increasing pressure on traditional nomadic practices. As Russian settlement intensified, bringing with it novel agricultural techniques and tools, clans played a crucial role in information dissemination and collective decision-making regarding their adoption—to their own detriment, as growing rates of sedentarization rendered their role increasingly obsolete, as I demonstrate in later sections.

2.3 Resettlement Administration, 1896–1918

In order to deal with increasing in-migrations of Russian and Ukrainian peasants into Central Asia after 1889, in 1896 the Tsarist government creates the Resettlement Administration within the Ministry of Internal Affairs. After the 1905 reforms, it was moved to the Main

[5]Though there are evidence that already prior to extensive Russian colonization winter stop lands have already began being partitioned into separate allotments over which nuclear families held more exclusive use rights

Administration of Land Management and Agriculture (GUZiZ), which later became the Ministry of Agriculture (Ananich and Ganelin, 2007). In 1906, peasant resettlement to the steppe became heavily politicized as the state began to be aggressively involved in its promotion. With the Resettlement Administration under the highly technocratic GUZiZ, land norms became central to the organization's "technocratic ethos," which posited that complete information could lead to the most efficient use of land and human capital (Kaufmann 1895, 1897). The technocratic ideology of the Resettlement Administration was rooted in a long-standing Russian political tradition of progressive statism. Officials in technical ministries like GUZiZ had developed an independent sense of identity and purpose, seeing themselves as "specialists" whose practical knowledge of applied sciences informed their formal roles. They believed in scientized state intervention and "productive" labor over "speculation," championing technocratic knowledge. They also saw the state's agenda as being at odds with the interests of the Russian nobility, who they considered "mere frictions." (Wortman 1976).

The Resettlement Administration was staffed by individuals deeply committed to the cause of scientifically managed colonization. Aleksandr Krivoshein, a key figure in GUZiZ, played a crucial role in restructuring peasant agriculture, first under Sergei Witte before 1905 and later under Petr Stolypin's reform efforts after 1906. His deputy, Aleksandr Rittich, shared his commitment to agricultural reform. Grigory Glinka, who headed the Resettlement Administration, was instrumental in training specialists in this specific area of state management. His deputy, Gennady Chirkin, prolifically advocated for settlement and land reform and co-edited the agency's semi-official journal, *Voprosy kolonizatsii*. Another influential official, Nikolai Gavrilov, was recognized for his planning and systematization skills. A younger generation of officials, like Aleksei Tatishchev, joined the Resettlement Administration drawn by its technocratic approach to social issues. Tatishchev, described as an "enthusiast," later served as field director in the Maritime Province and head of the Turkestan's Department of Agriculture. He, along with others, envisioned the colonization of the peripheries as a state-directed endeavor to maximize the human and productive resources of the empire as a whole (Tatishchev 2001, p. 250).

Despite its technocratic ambitions, the Resettlement Administration encountered numer-

ous challenges and contradictions in implementing its program, particularly in Central Asia. Officials sought to apply their model in Turkestan, expropriating lands claimed by the indigenous population for redistribution to Russian settlers. They calculated land "norms" based on abstract formulas and superficial investigations, often disregarding the existing land usage and nomadic lifestyle of the native Kazakh and Kyrgyz populations (Morrison 2012). The program faced criticism from officials like Konstantin Palen, who led a Senatorial investigative commission to Turkestan in 1908-1909. Palen criticized the Resettlement Administration's approach as blindly technocratic and divorced from reality, pointing out that their "norms," calculated in offices rather than on the ground, led to the expropriation of land without adequate compensation for the indigenous population. He argued that the resettlement officials' focus on maximizing productivity disregarded existing social and cultural dynamics, leading to administrative chaos and poverty in the region. The Resettlement Administration's program in Central Asia exposed the inherent contradiction in their technocratic approach. While they aimed to implement scientifically informed and efficient land management policies, their reliance on abstract "norms" and disregard for local contexts often led to unintended negative consequences, exacerbating social tensions and ultimately undermining their stated goal of achieving a "just fulfillment" of the colonization program.

2.4 Russian Colonization of the Steppe

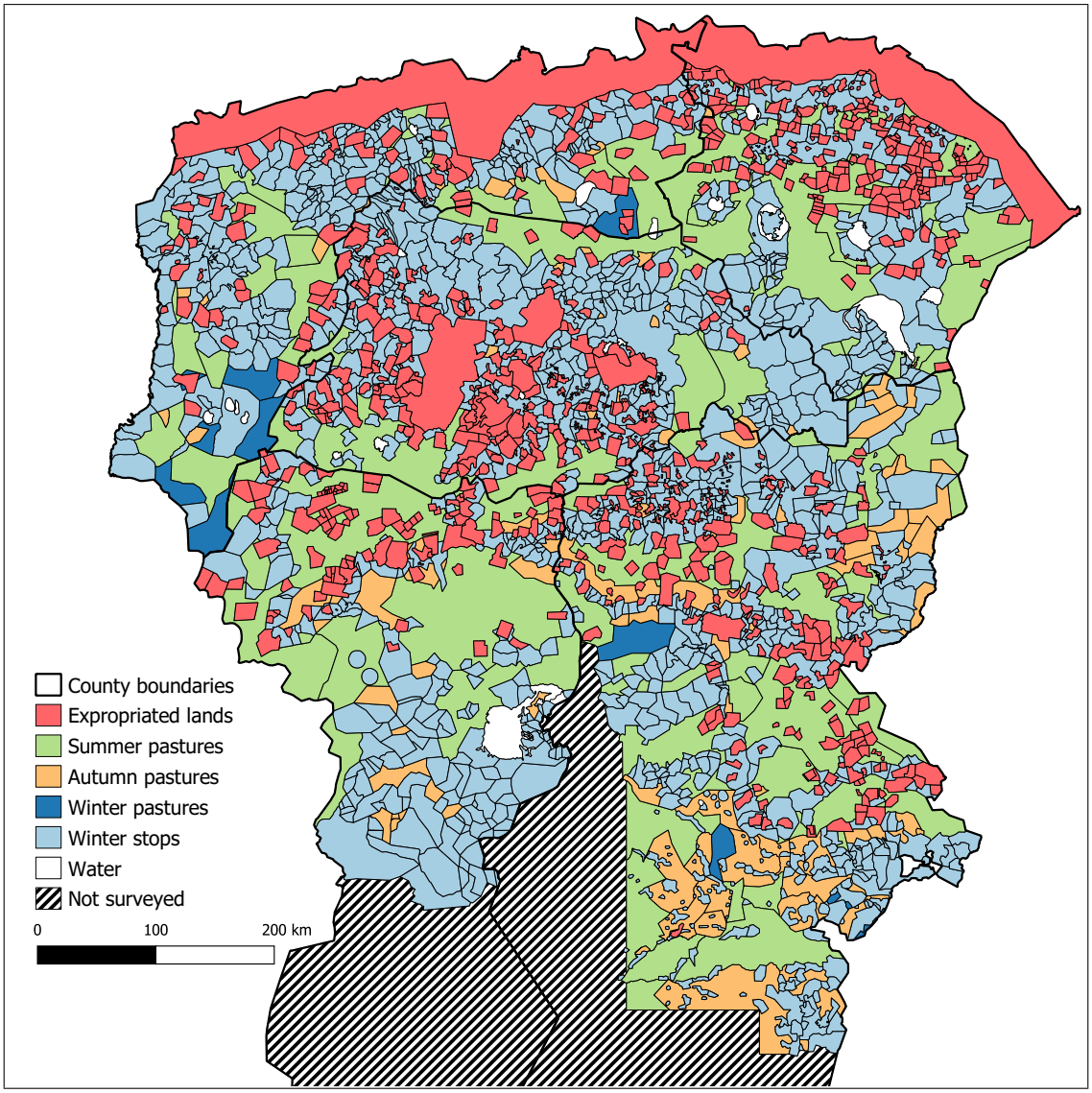
Kazakhstan became protectorate of the Russian Empire in the course of the first half of the 18th century, when Kazakh clans have been increasingly involved in military conflicts with their immediate neighbors—Oirats—over access to pastures and water sources in the north of modern Kazakhstan. Kazakhs of the Junior Horde became the first one to request such a status in 1731, followed by those of Middle and Senior Horde in 1735 and 1748, respectively.

While remaining a fairly autonomous entity through the 18th century, Kazakh Steppe was gradually transformed into a colony in the course of the 19th century via a series of military and administrative reforms. By 1822, Khan power was completely abolished in the Middle Horde, and in 1868 the multi-layered imperial system of territorial management was established from scratch, effectively coopting Kazakh landed elites and clan elders into newly established administrative roles, thus making Kazakh Steppe an integral part of the Empire

(Abuseitova et al. 2001, pp. 341-359).

Starting in the 17th century with runaway serfs escaping tax burden and oppression from their landlords, the initial migration from the core of the Empire into Central Asia was negligible, but it rapidly accelerated in the last quarter of the 19th century, reaching its peak in 1910s. One could roughly separate three phases in which migrations from the European core developed in the course of the 19th century. The first one lasted until approximately 1868 and is characterized by total dominance of the Cossacks and other military and paramilitary personnel among the early settlers. While granted wide autonomy in conducting their affairs with the indigenous population, they ultimately developed a certain type of secluded existence so typical for other Cossack communes elsewhere, and, therefore, had limited impact on the local economy. The second phase began after the emancipation of serfs in the Empire in 1861, which, among other things, untied millions of people from their land and allowed surplus labor to be syphoned elsewhere. These migrations were still limited and largely chaotic, yet the central government neither discouraged, nor encouraged such movements for two reasons: first, they correctly believed that should such necessity arise in the future, a coordinated colonization of Central Asia and Siberia would be a much easier endeavor if there is already a critical mass of loyal sedentary Russian population, and second, these migrations did somewhat ease the land pressure in the European part of Russia and in Ukraine.

Figure 2: The Extent of Russian Colonization of Akmolinsk Province by 1908



Note: This figure depict cadastral map of the surveyed part of Akmolinsk province from ca. 1908. Light blue polygons represent communes to which all 6340 extended households belong and where they spend most of their winters. I use a similar map from ca. 1897 and household's allocation across communes to calculate a number of geographic confounders at extended household level.

3 Data and Sources

I undertook a large-scale digitization of various archival and published primary sources to come up with a dataset featuring information on economic outcomes, nomadic genealogies, and potential confounders, as well as nomadic extended households' "eligibility" for land expropriations, combined with extremely granular data on actual land expropriations that took place between 1896 and 1918, retrieved from a number of archival and published sources produced by the local resettlement officials. The main data for this paper comes from statistical materials stemming from the two waves of colonial expeditions commissioned by the Resettlement Administration in 1896–1901 (headed by Fedor Shcherbina), which covered Akmolinsk and parts of Semipalatinsk and Turgai provinces, and in 1907-1913, which consisted of several independent expeditions and surveys in Turgai, Akmolinks (headed by Vasily Kuznetsov), Semirechye, Syr-Darya, and parts of Fergana, Samarkand, Semipalatinsk, and Uralsk provinces.

3.1 Household Data

For the sake of the exposition, this version of the paper only uses data from one province, Akmolinsk, due to relatively easier way one could match observations between the two waves of survey in this province. In particular, if a household was partitioned or merged with another household between the two waves, second-wave volumes for Akmolinsk province (Kuznetsov 1909, 1910*a*,*b*,*c*,*d*) allow to trace those changes manually and, therefore, to construct uniform synthetic households without the need to use fuzzy matching. After performing the matching procedure, I end up with 6340 unified extended households across the two waves in Akmolinsk province, which are distributed across its five counties in the following order: Atbasar (704), Akmolinsk (1617), Kokchetav (1847), Omsk (815), and Petropavlovsk (1357). In rare cases when parts of the same unified household end up belonging to different *zemel'nye obshchiny* (communes) in either of the waves, I assign smaller parts to a bigger part's commune.

Both Shcherbina's (1898, 1902*a*, 1902*b*, 1907, 1908) and Kuznetsov's expeditions typically took place during summer months, usually from early June to early September, when

nomadic families were staying at their summer pastures. Surveys were conducted by teams composed of Russian statisticians and land engineers and Russian-speaking Kazakh or Tatar interviewers. The resulting corpus of data constitute a detailed agricultural census: virtually every household that had a winter stop in one of Akmolinsk's province counties was covered by the survey, unless it, for whatever reason, spent last summer and a winter migrating with another clan outside of the province, in which case it would have been recorded as absent. The main questionnaire was at the extended family level and was typically answered by the heads of extended families (normally, the eldest or the richest male member, called *aqsaqal*, "the white-bearded man"). The questionnaire consisted of several sections: demographic and geographic variables (access to fresh water, name of the winter stops, names of nearby water sources, etc.), clan self-identification (name of the male ancestor all members of the extended household identify with), detailed sections on livestock and agricultural activities (pastoralism, cropmaking, haymaking, both on extensive and intensive margins), ownership of modern agricultural tools (ploughs, harrows, manual and horse-driven seeders, etc.), permanent buildings on winter stops, participation in contractual labor markets (number of hired workers on annual and half-annual contracts, number of members of households employed outside of the household) and in rental market for land.

The main goal of these expeditions was finding out how much surplus land was in the hands of the nomads and, therefore, how much land could be taken away for the needs of the resettlement movement without disrupting the nomadic economy, but with the ultimate aim of gently facilitating its transition to sedentarization. However, very few people were privy to the clandestine aims of these surveys: for example, Shcherbina himself was a strong believer in the fact that a limited peasant settlement would not disrupt the nomadic economy, yet already in 1900 he criticized the scale of expropriations and expressed his frustration about the way the Resettlement Administration used his expedition's results in developing the "norms" of nomadic land use (Shcherbina 1900).

Despite explicitly political motivation behind these surveys, both contemporary practitioners and later historians confirmed high quality and reliability of these data. For example, Kaufmann (1907) questioned Resettlement Administration claim about potential underdeclaration of livestock by nomadic families and proposed an adjustment to subse-

quently used norms that, seemingly, had no effect on Resettlement Administration's calculations as neither Kuznetsov's, nor census volumes covering other provinces contain any reference to Kaufmann's argument. Rumiantsev (1910), who have launched the expedition to Semirechye province a year after writing this report, while pointing out some inconsistencies in Shcherbina's classification of households by livestock wealth, praised the overall quality of collected data. A number of Soviet historians, including Shakhmatov (1964) and Tolybekov (1971), have demonstrated that the key economic outcomes recorded by the expeditions are in line with archaeological and anthropological data measuring the same socio-economic variables.

Overall, the volumes of *"Materialy po kirgizskomu zemlepol'zovaniiu..."* by Shcherbina, Kuznetsov, and others remain one of the best examples of high-quality statistical data produced by the Russian Czarist bureaucracy, on par with the corpus of 19th century local *zemstvo* statistics, early modern *pistsovye knigi* tax cadasters and the 18th century General Land Survey cadastral maps and accompanying economic commentaries, which only makes relative obscurity of *"Materialy..."* even more puzzling.

3.2 Nomadic Genealogies

Another important piece of information contained in the data appendices to *"Materialy..."* are detailed genealogical trees called *shezhire* which constitute the basis of Kazakh society. These trees link the top-level clan identifications (which, for the most part, are names of semi-legendary clan founders) to male ancestors that extended families self-report as being most identified with in the main questionnaires. I supplement genealogical information provided by the *"Materialy..."* volumes with more modern research on Kazakh ancestral kinship networks from Tynyshpaev (1925) and Vostrov and Mukanov (1968, 1974).

I was able to identify and partition all 6340 harmonized extended households residing in Akmolinsk province into five tribes and 27 clans of the Middle *jüz*—Argyn (*atygai, besentein, begendyk, qanzhygaly, qarakesek, qarauly, kuandyk, suyindyk, tarakty, tobykty, shegendyk*), Kerei (*balta, koshebe, siban, taryshy*), Naiman (*baltaly, baganaly, qöqzharly*), Qypshak (*butyn, qarabalyk, koldenen, kulan-qypshak, turaigyr, tory*), Uwaq (*barzhaqsy, bidaly, erenshi, zhansary, sarman, shaigez, shoga*),—as well as insignificant numbers of members

of Junior—Zhetyru (*zhagalbaily, kereit, ramadan, tabyn, tama, teleu*)—and Senior Hordes (Dulat, Kangly, Zhalayiryly), and members of "special" tribes which are not included in Kazakh genealogies, such as Töre, Kozha (or *Khoja*), and Kyrgyz. Overall, after getting rid of 118 households belonging to "special" tribes, I am able to partition the remaining 6222 harmonized households between 1029 nested lineages. Then, for each of these lineages, I am able to calculate generational distance between the clan's alleged founder (top of the lineage) and self-reported male ancestor a household identifies with.

3.3 Cadastral Maps and Geography

Each county-level volume of the *"Materialy..."* also contains a cadastral map of nomad land use by the time of the survey. I digitize all of these maps from both waves of the survey and use QGIS to georeference and vectorize the resulting raster images. Figure 2 presents the result of this procedure for Akmolinsk province. Polygons of different colors represent land used for winter stops, summer, spring, autumn, and special winter pastures, with lands confiscated by 1908 colored in red. Polygons colored in light blue represent winter stops, which includes high quality "insurance" lands in nomads posessions, closed-access high quality winter pastures, as well as closed-access water sources. Matching extended households from my sample to these polygons allows me to calculate several potential geographical confounders, such as average yearly precipitation in mm per year and FAO GAEZ alfalfa suitability index under low input and no artificial irrigation, i.e. under rainfed regime. I also construct another variable which measures the number of days per year an extended household has access to fresh water by combining information on household-level access to various water sources from *"Materialy..."*.

3.4 Measuring Expropriations

Unfortunately, the original *"Materialy..."* volumes do not provide systemic data regarding what households were expropriated and by how much. I therefore had to refer to archival sources in order to construct my treatment variable. Luckily, it appears that local resettlement officials, as well as the bureaucrats responsible for transportation of settlers to their

final destinations have been meticulously keeping track of all expropriated land in Central Asia, including information on what nomadic households were subject to confiscations and were, therefore, owned a compensation.

To that end, I digitize all existing published county- and province-level land expropriation reports produced by either the Resettlement Administration or by the local resettlement officials, survey parties, or transportation bureaucrats[6], and I supplement them with digitized contemporary primary sources on statistics of land management from archives in Chelyabinsk (Russia) and Almaty (Kazakhstan). Both CSA of Kazakhstan and OSA of Chelyabinsk Oblast contain annual plot confiscation reports which allow me to link each individual expropriated plot to extended households from the census that were affected by this particular expropriation[7]. Despite the fact that these expropriation reports contain plenty of plot-level data such as the size of the plot, quality of soils, proximity to water sources, there is no way to tell how exactly, say, the size of expropriated land was allocated among different households in cases when several were affected. I therefore only record a binary measure of land expropriations which I use as my treatment variable.

3.5 Agricultural Productivity

To measure the impact of land expropriations on crop productivity, I digitize 10 waves—from 1880 to 1912—of county heads' surveys conducted every three years or so to collect data on population, land use and crop productivity. Data on 1915 and 1917 comes from province-specific volumes studying agricultural development and the 1917 agricultural census tabulated at subcounty level. The resulting sample consists of 352 subcounties partitioned between 20 counties and 4 provinces. The two outcome variables are the amount of croplands per nomadic household used by the nomads in a subcounty to crop one of the five crops, and the cropland area-weighted average crop productivity across those five crops, where crop-specific productivity is measured as a ratio of yield to seeds used.

[6]For examples of such public reports, see Ministry of Agriculture and State Property. Department of State Landed Property, (1900), Ministry of the Interior. Resettlement Administration, (1903), Molodykh (1906), to name a few.

[7]CSA of Kazakhstan: FF. I-318, I-393, I-460, I-700, I-828; Open State Archive of Chelyabinsk Oblast: F. I-13, op. 1, dd. 209, 218, 395, 414, 434, 444, 446, 450, 489, 523, 586, 596, 809, 823, 853, 993.

4 Empirical Strategy

The impact of land reforms is challenging for economists to measure for several reasons. First, the high cost and large potential returns of such policies mean that few policymakers, especially in a 19th century colonial context, are willing to allow for arbitrary random treatment. Intentional targeting of large indigenous langholding elites, political favoritism, and rent-seeking potential would lead policy treatment to be correlated with other simultaneous government programs (such as, for example, construction of railroads and canals), connectedness to markets, and measures of economic growth potential, all of which would bias naive estimates in an unknown direction. Second, in order to measure the impact of a land reform on social structures, one needs a setting where a reform leads to a truly sweeping societal changes with large treatment samples. Sample post-reform surveys therefore are unlikely to have sufficient number of treated and control observations. In contrast, analysis at more aggregate level is underpowered and might be subject to even greater identification concerns. My paper addresses these issues by considering one of the most meticulously documented massive land-grab efforts: the late 19th and early 20th century settler-colonization of Central Asia by the Russian Empire. In particular, I combine quasirandom variation in pre-designed land expropriation rules with household-level census data of affected nomadic population conducted both before and after most expropriations took place.

I obtain causal identification from the land expropriation guidelines designed by the Russian colonial Resettlement Administration following the 1896-1901 survey, according to which households were deemed eligible for confiscations (unbeknownst to them) based on the quality-weighted amount of land they de facto used for their economic activities, such as pastoralism, crop- and haymaking. As previously described, extended nomadic households in which an average nuclear family possessed an amount of land which allowed maintenance of more than 24 horses through a typical winter under pure nomadic pastoralism (i.e. that one which assumes extensive animal gazing with little to no stored fodder) were deemed eligible for expropriation. A final selection into treatment (i.e. land expropriation) was a complicated legal and logistical matter involving land engineers, county heads, local resettlement officials, leaders of the indigenous community, and arriving settlers, so it might have been partly

Table 1: Summary Statistics and Balance

	Full sample (1)	Below thres-hold (2)	Above thres-hold (3)	Diff. in means (4)	p-val. on diff. (5)	RD est. (6)	p-val. on RD est. (7)
Panel A: Δ_t Outcomes							
Self-ID (generations)	3.409	3.285	3.543	0.258	0.000	0.265	0.022
Russian speakers (share)	0.201	0.053	0.361	0.308	0.000	0.485	0.000
Rental land (*desyatinas*)	0.394	0.334	0.460	0.126	0.000	0.204	0.005
Hired workers (num.)	2.866	2.361	3.411	1.049	0.000	0.875	0.069
Agricultural tools (num.)	2.249	1.820	2.713	0.892	0.000	0.728	0.047
Permanent buildings (num.)	0.969	0.784	1.169	0.385	0.000	0.261	0.111
Cropland (*desyatinas*)	1.999	1.638	2.389	0.752	0.000	0.341	0.127
Hayland (*desyatinas*)	1.398	1.140	1.676	0.537	0.000	-0.047	0.870
Livestock (num.)	-1.094	-0.889	-1.315	-0.426	0.000	-0.388	0.000
Panel B: Confounders							
Alfalfa suitability (index)	3941.57	3955.50	3926.54	-28.969	0.440	117.13	0.388
Precipitation (mm)	262.107	262.742	261.421	-1.321	0.392	0.952	0.159
Access to fresh water (days)	273.322	273.428	273.206	-0.222	0.868	-5.556	0.275
Observations	6,340	3,291	3,049				

Note: The table presents mean values for extended households' outcomes (Panel A) and other characteristics (Panel B), measured either as a difference between post-treatment period ca. 1908 and pre-treatment period ca. 1897. All outcome variables come from household-level data appendices from the *"Materialy..."* volumes, while the three confounders come from FAO GAEZ (alfalfa suitability index under low input and rainfed regime and average annual precipitation in mm) and from *"Materialy..."* volumes (access to fresh water in days per year). Columns 1-3 show the unconditional means for all households, households below the treatment threshold, and households above the treatment threshold, respectively. Column 4 shows the difference of means across columns 2 and 3, and column 5 shows the p-value for the difference of means. Column 6 shows the regression discontinuity estimate, following the main estimating equation, of the effect of being above the treatment threshold on the baseline variable, and column 7 is the p-value for this estimate, using heteroskedasticity robust standard errors. An optimal bandwidth of ± 7 around the 24-horses thresholds has been used to define the sample of extended households (see text for details).

determined by political favoritism or various bureaucratic bottlenecks, but as long as the initial expropriation rule was followed to any degree, the likelihood of treatment should discontinuously increase at 24-horses threshold, which allows me to estimate the effect of land expropriations using a fuzzy regression discontinuity design.

As I noted before, in the current version of the paper, the sample consists of 6,340 harmonized extended nomadic households residing in all five counties of Akmolinsk province of the General-Government of the Steppe of the Russian Empire, of which 3,049 were deemed eligible for expropriations, yet only 1,464 actually experienced confiscations, including some households that were not deemed suitable for expropriation by Shcherbina's survey but were expropriated nonetheless for reasons discussed above.

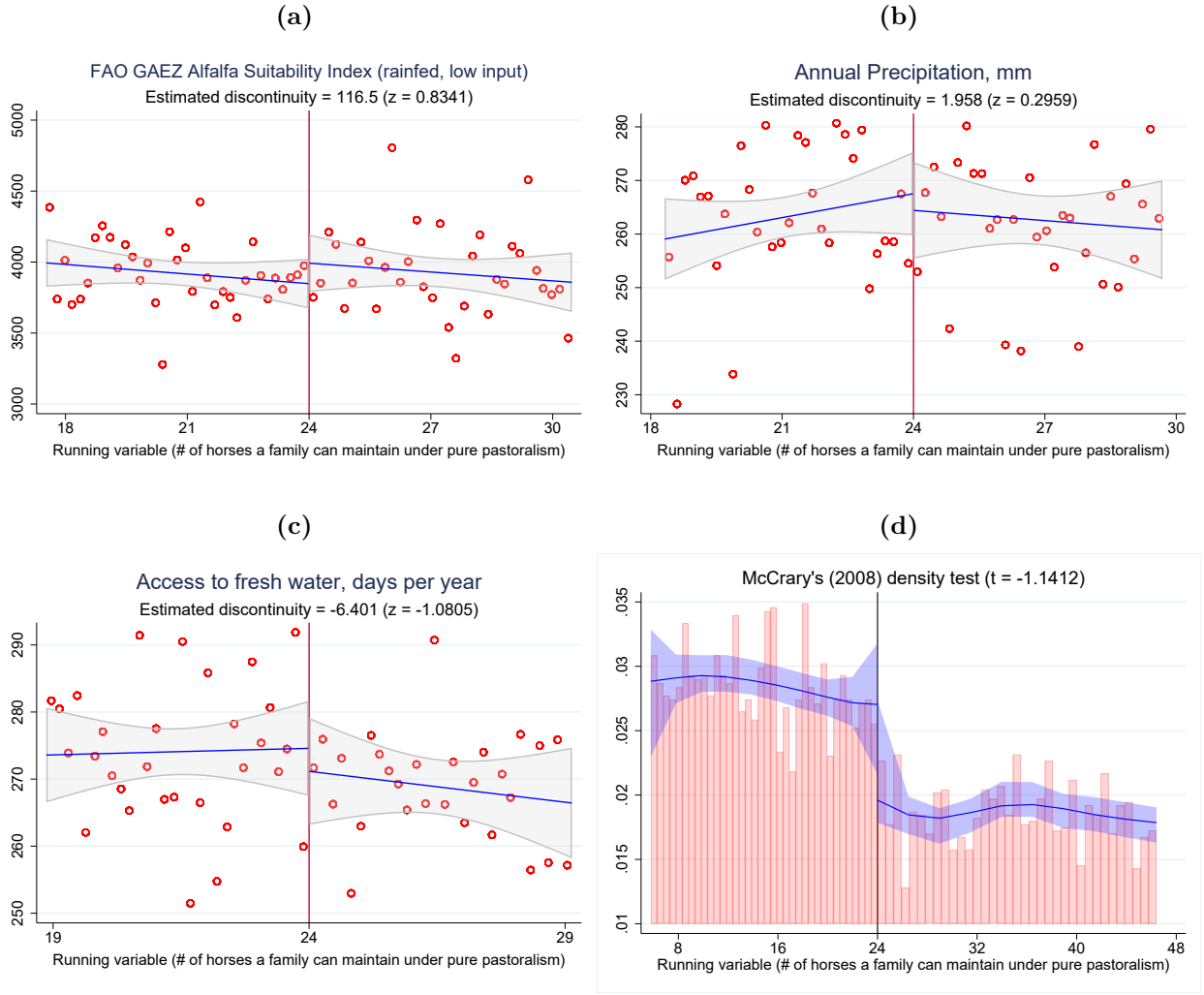
Assuming that all other potential household-level confounders run smoothly along the treatment threshold, the fuzzy RD estimator measures the local average treatment effect (LATE) of land expropriations for a household of land holdings able to support 24 horses, that is, for a household located exactly at the threshold. Following advice from Imbens & Lemieux (2008), my main RD specification uses local linear regression within a given optimal bandwidth on either side of the treatment threshold. Specifically, I use the following two-stage instrumental variable specification:

$$\Delta_{1908-1897} y_{h,d} = \beta_0 + \beta_1 Expropriated_{h,d} + \beta_2 X_{h,d} + f(H_{h,d}) + \mu_d + \varepsilon_{h,d}, \tag{1}$$

$$Expropriated_{h,d} = \gamma_0 + \gamma_1 1(H_{h,d} \geq 24) + \gamma_2 X_{h,d} + f(H_{h,d}) + \eta_d + \upsilon_{h,d}. \tag{2}$$

Here, $\Delta_{1908-1897} y_{h,d}$ is the difference in levels of the outcome of interest in household h and survey district d between 1908 and 1897, $H_{h,d}$ is the number of adult horses an average nuclear family of extended household h in district d can support under pure pastoralism given the quality-weighted amount of land they have access to; $X_{h,d}$ is a vector of household-level controls measured in ca. 1897, which includes average annual precipitation (in mm) and alfalfa suitability index under low input and rainfed regime from FAO GAEZ, as well as the typical number of days in a year a household has access to fresh water from *"Materialy..."*; $f(H_{h,d})$ is the RD polynomial which controls for a smooth function of total land holdings of a household, expressed in adult horses these lands can sustain, and μ_d and η_d are survey

Figure 3: Balance of Household Characteristics & McCrary's (2008) Density Test



Note: The first three figures (**a**, **b** and **c**) plot ca. 1897 household characteristics over the running variable—the number of adult horses an average nuclear family within an extended household can maintain under pure pastoralism. Number of bins is selected by mimicking variance evenly-spaced method using spacing estimators. Regressions are estimated using local linear polynomials in the number of horses a family can potentially maintain under pure pastoralism, separately on each side of policy threshold of 24 horses on a sample of MSE-optimal bandwidth. Standard errors are clustered at the survey district level.
The last figure (**d**) shows the distribution of household land holdings expressed in the number of adult horses they are able to maintain under pure pastoralism around the treatment threshold of 24 horses. The blue line represents a quadratic polynomial fit to each half of the distribution following McCrary (2008), testing for a discontinuity at 24. The point estimate for the discontinuity is -0.01, with $p-$value of 0.2538.

district fixed effects. Across all specifications using household-level data, standard errors are clustered at survey district level.

Regression discontinuity estimates can be interpreted causally if covariates and the density of the running variable are balanced across the treatment threshold. Table 1 presents the mean values for the first difference of the key household-level outcomes studied in this paper (Panel A), as well as the set of geographical controls used in all regressions henceforth (Panel B). I confirm that there are no significant differences in potential geographic confounders between households above and below the treatment threshold. Figures 3a, 3b and 3c show the graphical version of the balance test for geographic confounders—all three of them turn out to be continuous at the treatment threshold. Figure 3d demonstrates that the density of the running variable is also continuous across the treatment threshold: the McCrary test statistic is -1.14, which corresponds to a p-value of 0.25[8].

Figure 4 shows the share of households which experienced land expropriations before ca. 1908 relative to the treatment threshold. There is substantial discontinuous increase in the probability of treatment at the threshold. Table 2 presents first stage estimates using the main estimating equation at various bandwidths. Crossing the treatment threshold raises the probability of being treated by 23-30.3 percentage points, and the estimates are robust to bandwidth choice.

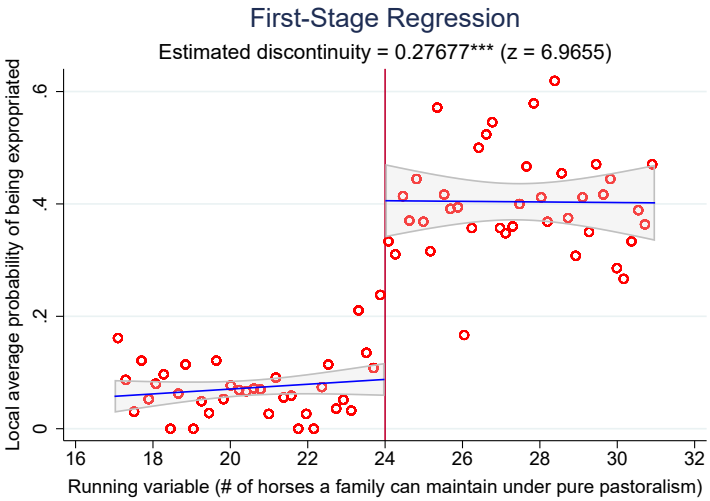
Finally, to flexibly estimate the effect of land expropriations on agricultural productivity, I use *volost*-level data on nomadic population, size of cropland by five different crops, and crop-level productivity and estimate the following simple event-study model:

$$y_{v,t} = \mu_v + \eta_t + \sum_{\tau=-q}^{-1} \gamma_\tau D_{v,\tau} + \sum_{\tau=0}^{m} \delta_\tau D_{v,\tau} + X_v + \varepsilon_{v,t}, \tag{3}$$

Here, μ_v is *volost* fixed effects, eta_t is survey year fixed effects, $D_{v,t}$ is a binary treatment variable, and X_v is a vector of time-invariant geographical confounders.

[8]I argue that sorting across the threshold was impossible due to the timing land use norms for nomads were publicized. Land on nomadic winter stops and climate were studied by land engineers independently from the nomads who at the time were oscillating between their summer and autumn pastures. Despite the fact that the Resettlement Administration started developing the norms for expropriation shortly before the survey in Kokchetav county was over by mid-September of 1896, the final formulas allowing to reverse-engineer the norms weren't publicized until 1903, when the first data volume was published, by which time confiscations were already underway.

Figure 4: First Stage: Effect of Expropriation Eligibility on Expropriation



Note: The figure plots the probability of getting expropriated by ca. 1908 against a number of horses an average nuclear family from a household can maintain under pure pastoralism regime. The sample consists of 2,035 households within a bandwidth of 7 of the 24-horses threshold.

Table 2: First Stage: Effect of Expropriation Eligibility on Expropriation

	±5	±6	±7	±8	±9	±10
Expropriation Eligibility	0.235	0.257	0.277	0.287	0.297	0.303
	(0.047)	(0.043)	(0.040)	(0.037)	(0.035)	(0.033)
Geo Confounders	Y	Y	Y	Y	Y	Y
Survey District FE	Y	Y	Y	Y	Y	Y
F-statistic	297.1	355.6	391.3	465.2	536.46	610.27
Observations	1,459	1,768	2,035	2,275	2,634	2,953
R^2	0.19	0.19	0.19	0.20	0.20	0.20

Note: This table presents first-stage estimates of the effect of being above the treatment threshold on a household's probability of treatment. The dependent variable is a indicator variable that takes on the value 1 if a household has experienced land expropriation prior to ca. 1908. The first column presents results for households with land holdings that are able to support from 19 to 29 adult horses under pure pastoralist regime, i.e., within 5-horses range around the 24-horses cutoff threshold. The second through sixth columns expand the sample to include households with land holdings able to support 6-, 7-, 8-, 9, and 10 more or less horses compared to the 24-horses cutoff threshold. The specification includes all geographical confounders presented in Table 1, as well as survey district fixed effects. Standard errors clustered at the survey district level are reported below point estimates.

5 Results

I start by presenting the main fuzzy regression discontinuity estimates of the impact of land expropriations on a number of key household-level outcomes expressed as a change between ca. 1897 and ca. 1908. Table 3 shows the results of these exercises. An exposure to land grab before ca. 1908 speeds up transition to sedentary agriculture and deterioration of traditional social structures that governed pastoralist activities. As such, extended households that experienced expropriations report going almost entire generation further down the genealogical tree between ca. 1897 and ca. 1908 when asked about their self-identification. This difference reflects their faster transition to a mode of production where economic activities are concentrated at a much tighter social circle, namely at extended household or a nuclear family level. Expropriated households also report faster rates of adoption of the Russian language by adult members of the households. The difference between the two groups is 1.63 percentage points, compared to 0.05 percentage points mean increase for the control group. Figure 5 shows reduced form estimates of the effect of expropriations on the same set of outcomes. Note again that the table point estimates are larger than the jumps observed in the figures because the tables present fuzzy RD (IV) estimates, while the figures show the reduced for difference at the 24-horses threshold.

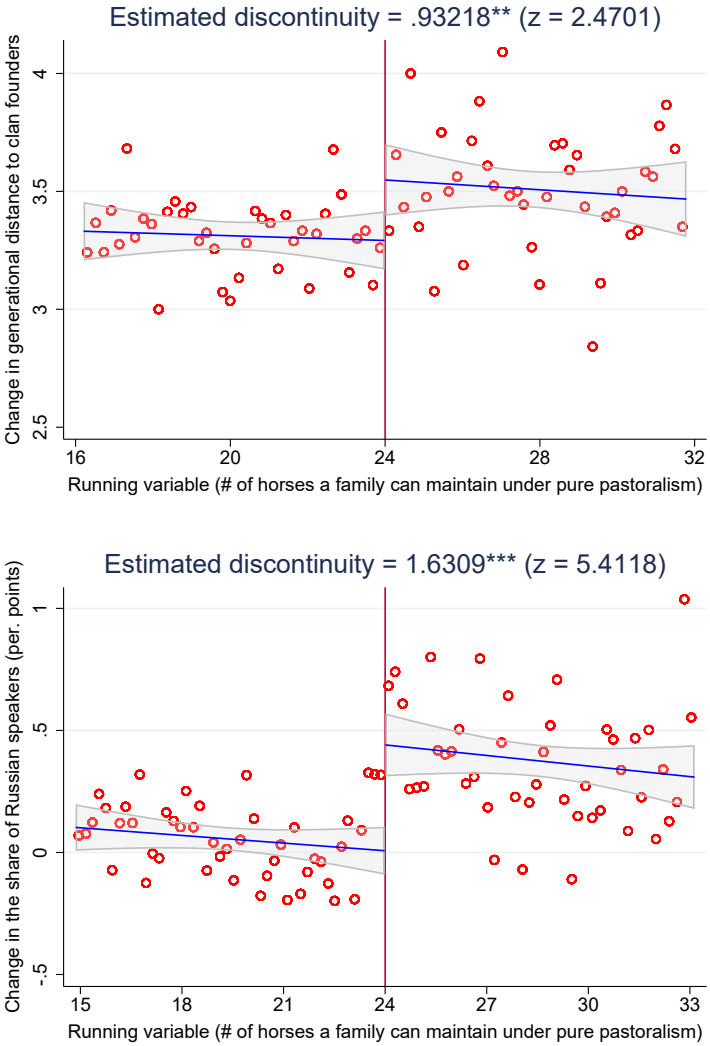
Next group of outcomes measures households' exposure to emerging markets for hired labor and rental land. Expropriated household report hiring and sending out a whooping 4.4 workers per nuclear family more than unaffected households. They also increase their cumulative exposure to the two-sided rental market of land by almost an entire *desyatina* (2.3 acres). Both of these results reflect the fact that expropriated households were forced to rapidly switch to a more intensive, sedentary mode of production. As before, Figure 6 presents corresponding graphical reduced form estimates. Next, I estimate the impact of expropriation on the per-family size of livestock herd, expressed in the number of adult horses. One of the reasons of maintaining larger herds is to account for relatively higher livestock mortality due to adverse weather under pure pastoralist regime when the nomads typically do not make any special arrangements for producing fodder beforehand or during the winter and just let their herds gaze on winter pastures. Once transition to more intensive

Table 3: Impact of Expropriations on Household-Level Outcomes

	Culture		Markets		
	Self-ID *generations*	Russian Speakers *share*	Hired Workers *count*	Rental Land *desyatinas*	Livestock *count*
	(1)	(2)	(3)	(4)	(5)
Expropriation	0.932	1.631	4.384	0.790	-1.451
	(0.377)	(0.301)	(2.182)	(0.307)	(0.427)
Geo Confounders	Y	Y	Y	Y	Y
Survey District FE	Y	Y	Y	Y	Y
Control Group Mean	3.311	0.056	2.541	0.316	-0.889
Bandwidth	7.795	9.144	5.549	6.327	6.551
Eff. Observations	2,243	2,688	1,604	1,865	1,928
R^2	0.26	0.28	0.22	0.24	0.25
	Investment		Land Use		
	Agro Tools *count*	Permanent Buildings *count*	Crop Land *desyatinas*	Hay Land *desyatinas*	—
	(6)	(7)	(8)	(9)	—
Expropriation	2.628	0.942	1.261	-0.440	—
	(1.335)	(0.539)	(0.724)	(1.164)	—
Geo Confounders	Y	Y	Y	Y	—
Survey District FE	Y	Y	Y	Y	—
Control Group Mean	1.888	0.774	1.636	1.208	
Bandwidth	6.959	6.948	8.484	6.212	—
Eff. Observations	2,025	2,021	2,465	1,835	—
R^2	0.25	0.25	0.27	0.24	—

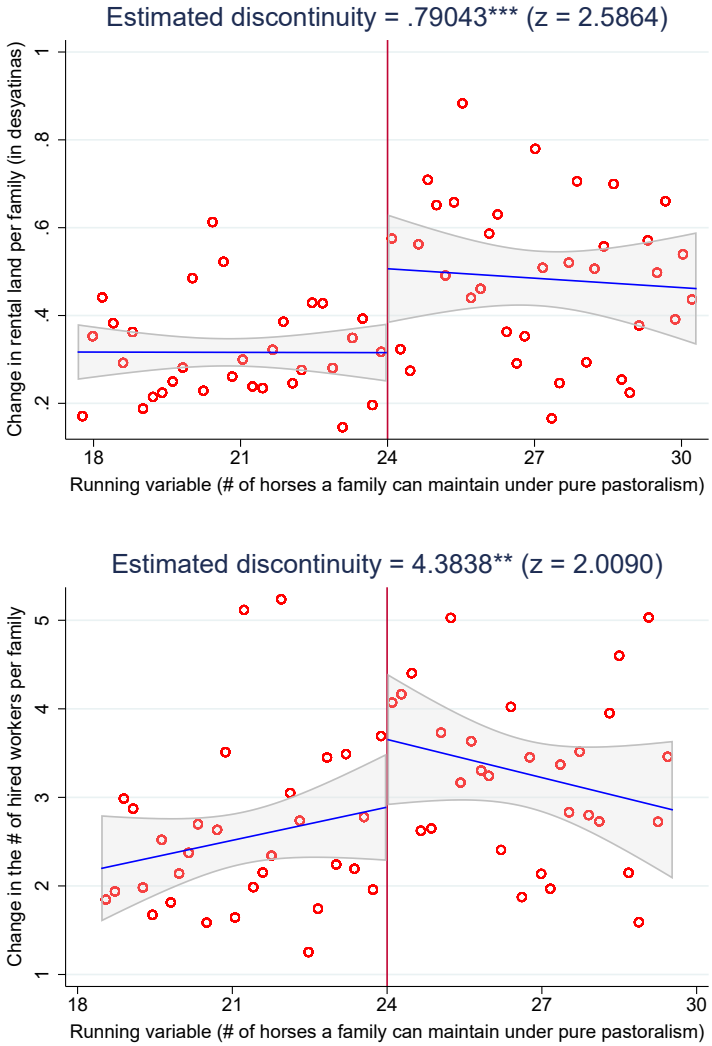
Note: This table presents fuzzy regression discontinuity estimates from the main estimating equation of the effect of land expropriation on changes in social structures and culture, exposure to markets for hired labor and rental land, investments in new agricultural tools and construction of permanent buildings, extensive margin of land use, and units of livestock between ca. 1897 and ca. 1908. Regressions are estimated using local linear polynomials in the number of horses a family can potentially maintain under pure pastoralism, separately on each side of policy threshold of 24 on a sample of MSE-optimal bandwidth. For each regression, the outcome mean for the control group (households with running variable values below the threshold) is also shown. The specification includes baseline household-level geographic controls, as well as survey district fixed effects. Standard errors clustered at the survey district level are reported below point estimates.

Figure 5: Reduced Form: Impact of Expropriations on Self-Identification and Share of Russian Speakers



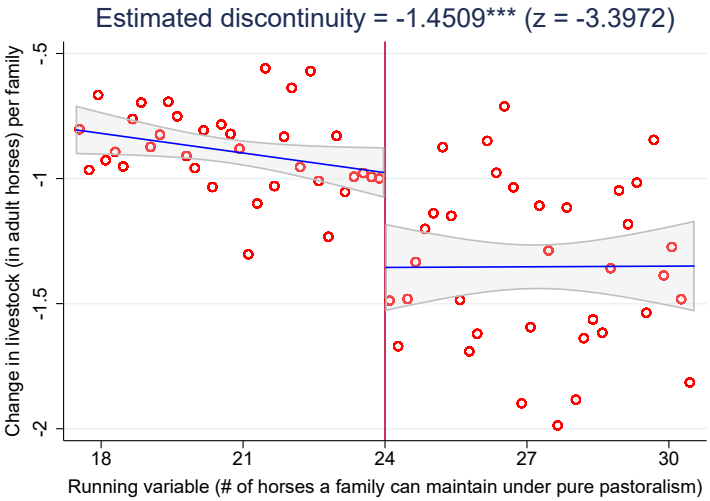
Note: The figures present the reduced form RD plots of the impact of land expropriations on the change in household's self-reported generational proximity to their clan's founders (top) and the change in the share of Russian speakers between ca. 1897 and ca. 1908 (bottom). Number of bins is selected by mimicking variance evenly-spaced method using spacing estimators. Regressions are estimated using local linear polynomials in the number of horses a family can potentially maintain under pure pastoralism, separately on each side of policy threshold of 24 horses on a sample of MSE-optimal bandwidth. Standard errors are clustered at the survey district level.

Figure 6: Reduced Form: Impact of Expropriations on Rental Land and Hired Workers



Note: The figures present the reduced form RD plots of the impact of land expropriations on the change in the amount of rental land expressed in *desyatinas* (1 *desyatina* = 1.09 ha = 2.7 acres) and the change in the number of hired workers per family between ca. 1897 and ca. 1908. Number of bins is selected by mimicking variance evenly-spaced method using spacing estimators. Regressions are estimated using local linear polynomials in the number of horses a family can potentially maintain under pure pastoralism, separately on each side of policy threshold of 24 horses on a sample of MSE-optimal bandwidth. Standard errors are clustered at the survey district level.

Figure 7: Reduced Form: Impact of Expropriations on Livestock Ownership



Note: The figure presents the reduced form RD plots of the impact of land expropriations on the change in the number of livestock units expressed in adult horses per family between ca. 1897 and ca. 1908. Number of bins is selected by mimicking variance evenly-spaced method using spacing estimators. Regressions are estimated using local linear polynomials in the number of horses a family can potentially maintain under pure pastoralism, separately on each side of policy threshold of 24 horses on a sample of MSE-optimal bandwidth. Standard errors are clustered at the survey district level.

agriculture is underway, one no longer needs to maintain large herds because the weather-related mortality risk could be alleviated by winter hay making or storing hay, straw or other types of fodder procured in the fall throughout the winter.

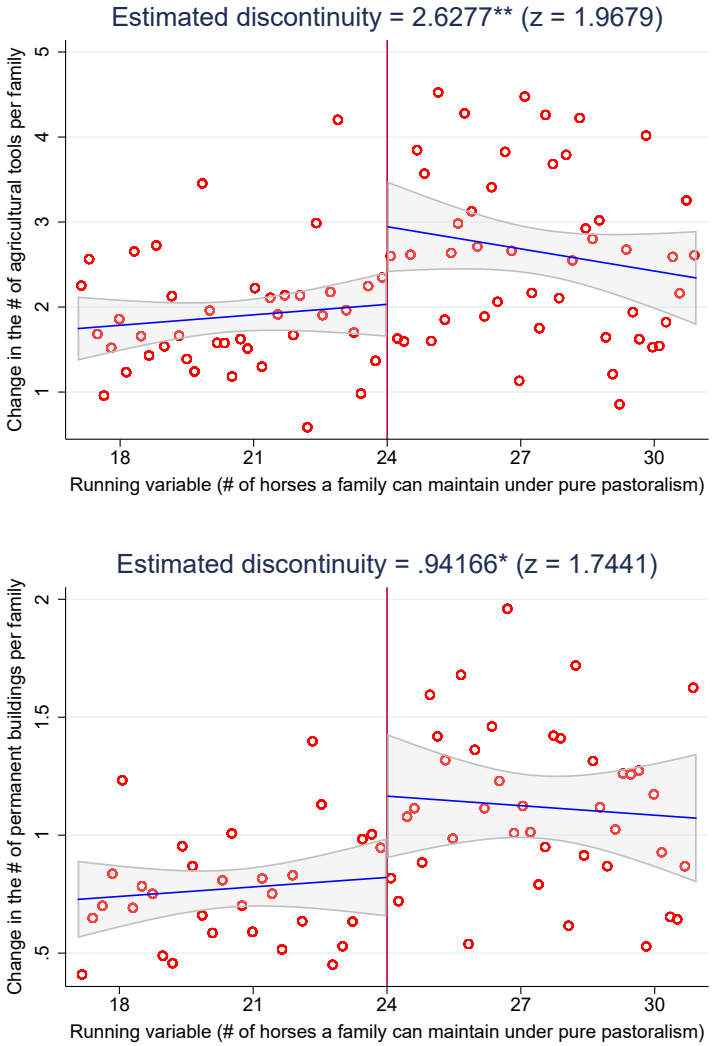
Next set of outcomes comprises investments into purchasing or producing agricultural tools like ploughs, manual or horse-driven seeders, rakes, etc. and construction of permanent buildings, both utility and residential. Columns (6) and (7) show that expropriated households end up having 2.6 more tools per nuclear family and an entire permanent building more vis-a-vis unexpropriated households. Again, both results are in line with these households' more rapid transition to semi-sedentary mode of production where fewer members of the household participate in seasonal transhumant migrations and instead stay behind and work the lands on winter stops and elsewhere.

Columns (8) and (9) of Table 3 present the final set of RD estimates related to changes in extensive use of lands in response to land expropriations. There are evidence (although

point estimate is noisy) that expropriated households to increase per-family amount of land allocated to individualized cropmaking. As for making hay, it seems like there are no difference between treatment in control groups, suggesting that communal form of procuring hay using more marginal, commonly used pastures remained, at least for the time being, the dominant form of haymaking. Figure 9 demonstrates corresponding reduced form RD estimates in graphical form.

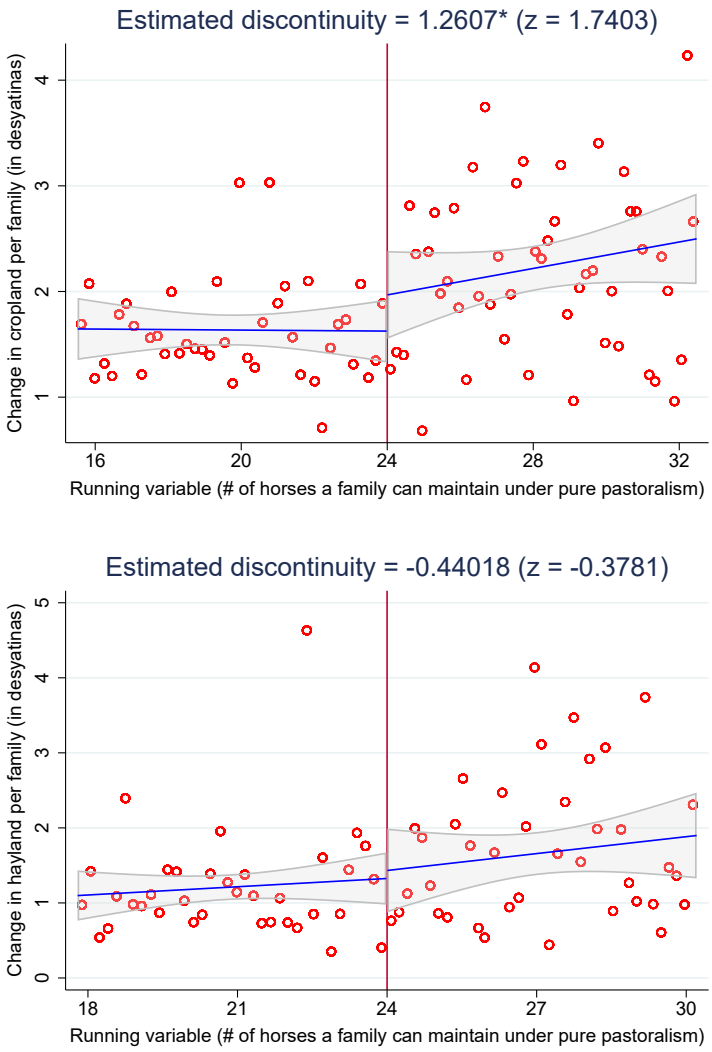
Finally, to supplement the RD estimates which use household-level data, I present event study estimates of the impact of expropriations on cropland (extensive margin) per family and area-weighted agricultural productivity (intensive margin) at a subcounty (*volost*) level. Figure 10 contains two event study plots where I control for subcounty and wave/year fixed effects, and cluster standard errors at a subcounty level. We see that land expropriations induced affected nomadic population to allocated larger portion of their lands towards cropmaking due to the fact that losing some of the best winter stop lands to settlers made pure nomadic lifestyle unsustainable, so to maintain their livestock they now have to switch parts of the lands from extensive animal gazing to crop- and haymaking. The bottom panel shows that despite some transfer of agricultural knowledge that might have happened as a result of peasant in-migrations, the fact that nomads were pushed into using more marginal pasture lands—which might have been more suitable for less productive crops—more intensively had a net negative and statistically significant impact on their crop productivity.

Figure 8: Reduced Form: Impact of Expropriations on Agricultural Tools and Permanent Buildings



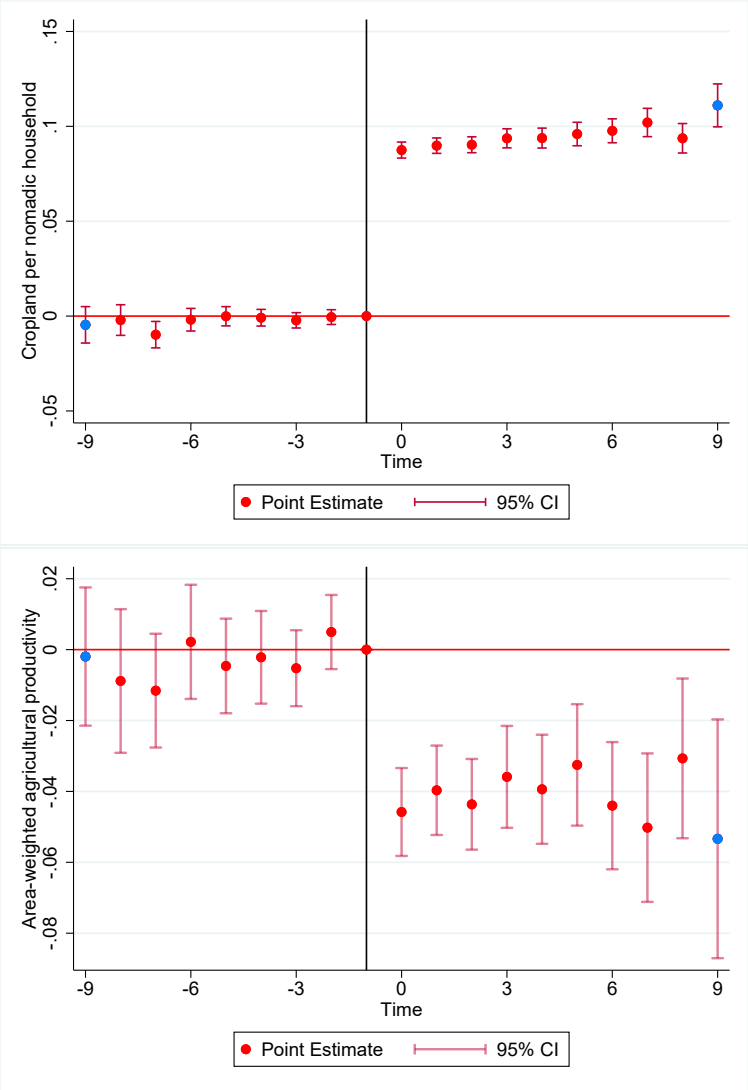
Note: The figures present the reduced form RD plots of the impact of land expropriations on the change in the number of agricultural tools (top) and permanent buildings (residential or utility, bottom) per family between ca. 1897 and ca. 1908. Number of bins is selected by mimicking variance evenly-spaced method using spacing estimators. Regressions are estimated using local linear polynomials in the number of horses a family can potentially maintain under pure pastoralism, separately on each side of policy threshold of 24 horses on a sample of MSE-optimal bandwidth. Standard errors are clustered at the survey district level.

Figure 9: Reduced Form: Impact of Expropriations on Croplands and Haylands



Note: The figures present the reduced form RD plots of the impact of land expropriations on the change in the amount of cropland (top) and hayland (bottom) per family between ca. 1897 and ca. 1908. Number of bins is selected by mimicking variance evenly-spaced method using spacing estimators. Regressions are estimated using local linear polynomials in the number of horses a family can potentially maintain under pure pastoralism, separately on each side of policy threshold of 24 horses on a sample of MSE-optimal bandwidth. Standard errors are clustered at the survey district level.

Figure 10: Event Study: Impact of Expropriations on Croplands and Agricultural Productivity



Note: The figures present event study plots of the impact of land expropriations on the amount of land devoted to growing 5 different crops (top) and area-weighted crop productivity across the same 5 crops (bottom). Standard errors are clustered at a subcounty (*volost*) level.

6 Conclusion

Throughout history, settler-colonialism is oftentimes associated with massive changes in the constraints and endowments faced by indigenous populations. In this paper, I have show that traditional social structures such as kinship networks determined by genealogical proximity play important role in adaptation to such changes.

Not only didn't these pre-existing social structures hinder adjustments to large-scale land expropriations experienced by Central Asian nomads, but it is likely that their built-in flexibility allowed the nomads to switch to semi-sedentary mode of production more smoothly. Prior to the active phase of colonization, clans and tribes were collective action vehicles that governed massive transhumant migrations of related households across seasonal pastures and regulated their collective use. As some households were experiencing increased land pressure due to peasant settlement and were, therefore, forced to sedentarize and start using the remaining lands in their possession more intensely, the role of clans has gradually faded away while the role of more tighter relationships—where most economic activities were now being concentrated—has increased, resulting in prompting formerly nomadic families to make economic decisions which weren't previously typical for them, such as investing in construction of permanent buildings, procuring fodder for animals instead of allowing them to gaze extensively throughout the year, which, in turn, allowed families to move away from precautionary and expensive herd-hoarding towards maintaining more balanced amount of livestock. I also demonstrate that as households were moving away from traditional institutions, the role of impersonal markets for labor and land became more salient, contractual wage labor became more widespread and some households became actively involved in renting lands to and from their immediate neighbours.

An important avenue for future research would be to see which aspects of clans remain important for society and national politics. On the surface, most Kazakhs are somewhat aware of their ancestry, yet it is unclear to what extend (if any) belonging to the same clan or tribe would facilitate interpersonal or intergroup relationships.

Beyond the spotlight on the social structures, my paper contributes to the debate on the role of pre-colonial institutions in long-run development, as well as to the literature studying

the effect of changes in resource scarcity or endowments on the well-being of indigenous populations. Finally, this is, I believe, one of the few studies of economic history of Central Asia and colonial policies of the Russian Empire.

References

Abuseitova, M. K., Abylkhozhyn, Z. B., Kliashtornyi, S. G., Masanov, N. E., Sultanov, T. I. & Khazanov, A. M. (2001), *Istoriia Kazakhstana i Tsentral'noi Azii*, Bylym, Almaty.

Acemoglu, D., Johnson, S. & Robinson, J. A. (2001), 'The Colonial Origins of Comparative Development: An Empirical Investigation', *American Economic Review* **91**(5), 1369–1401.

Akmolinsk Province (1874, 1887, 1908-1909, 1912-1916), *Pamiatnaia knizhka Akmolinskoi oblasti [i adres-kalendar] na... god*, Akmolinsk Province Board, Omsk.

Aldashev, G. & Guirkinger, C. (2017), 'Colonization and Changing Social Structure: Evidence from Kazakhstan', *Journal of Development Economics* **127**, 413–430.

Bardhan, P., Luca, M., Mookherjee, D. & Pino, F. (2014), 'Evolution of Land Distribution in West Bengal 1967–2004: Role of Land Reform and Demographic Changes', *Journal of Development Economics* **110**, 171–190.

Coleman, J. S. (1988), 'Social Capital in the Creation of Human Capital', *American Journal of Sociology* **94**, S95–S120.

De la Croix, D., Doepke, M. & Mokyr, J. (2018), 'Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Preindustrial Economy', *The Quarterly Journal of Economics* **133**(1), 1–70.

Dell, M. (2010), 'The Persistent Effects of Peru's Mining Mita', *Econometrica* **78**(6), 1863–1903.

Ensminger, J. (1997), 'Transaction Costs and Islam: Explaining Conversion in Africa', *Journal of Institutional and Theoretical Economics* pp. 4–29.

Fafchamps, M. & Lund, S. (2003), 'Risk-Sharing Networks in Rural Philippines', *Journal of Development Economics* **71**(2), 261–287.

Foster, A. D. & Rosenzweig, M. R. (2002), 'Household Division and Rural Economic Growth', *The Review of Economic Studies* **69**(4), 839–869.

Foster, A. D. & Rosenzweig, M. R. (2010), 'Microeconomics of Technology Adoption', *Annual Review of Economics* **2**(1), 395–424.

Governor of Akmolinsk Province (1870–1915), *Obzor Akmolinskoi oblasti za... god*, Akmolinsk Province Statistical Committee, Omsk.

Granovetter, M. (2005), 'The Impact of Social Structure on Economic Outcomes', *Journal of Economic Perspectives* **19**(1), 33—-50.

Greif, A. (1989), 'Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders', *The Journal of Economic History* **49**(4), 857–882.

Greif, A. (1993), 'Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition', *The American Economic Review* pp. 525–548.

Greif, A. (1994), 'Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies', *Journal of Political Economy* **102**(5), 912–950.

Greif, A. & Tabellini, G. (2017), 'The Clan and the Corporation: Sustaining Cooperation in China and Europe', *Journal of Comparative Economics* **45**(1), 1–35.

Guirkinger, C. & Aldashev, G. (2016), 'Clans and Ploughs: Traditional Institutions and Production Decisions of Kazakhs under Russian Colonial Settlement', *The Journal of Economic History* **76**(1), 76–108.

Guirkinger, C. & Platteau, J.-P. (2015), 'Transformation of the Family Farm under Rising Land Pressure: A Theoretical Essay', *Journal of Comparative Economics* **43**(1), 112–137.

Guiso, L., Sapienza, P. & Zingales, L. (2016), 'Long-term Persistence', *Journal of the European Economic Association* **14**(6), 1401–1436.

Gupta, B. (2014), 'Discrimination or Social Networks? Industrial Investment in Colonial India', *The Journal of Economic History* **74**(1), 141–168.

Imbens, G. W. & Lemieux, T. (2008), 'Regression Discontinuity Designs: A Guide to Practice', *Journal of Econometrics* **142**(2), 615–635.

Kaufmann, A. A. (1895), 'Khoziaistvennoe polozhenie pereselentsev, vodvorennykh na kazennykh zemliakh tomskoi guberniipo dannym proizvedennogo v 1894 g., po porucheniiu g. tomskogo gubernatora, podvornogo issledovaniia , 2 vols.'.

Kaufmann, A. A. (1897), *Materialy po voprosu ob organizatsii rabot po obrazovaniiu pereselencheskikh uchastkov. Iz otcheta starshego proizvoditelia rabot Kaufmanna po komandirovkie v Akmolinskuiu oblast lietom 1897 g.*, Ministry of Agriculture and State Property. Department of State Landed Property, St. Petersburg.

Kaufmann, A. A. (1907), *K voprosu o proiskhozhdenii russkoi zemel'noi obshchiny*, Moscow.

Konshin, N. (1896), *Kratkii statisticheskii ocherk promyshlennosti i torgovli v Akmolinskoi oblasti za 1880-1894*, Akmolinsk Province Statistical Committee, Omsk.

Kuznetsov, V. K. (1909), *Kirgizskoe khoziaistvo v Akmolinskoi oblasti.*, Vol. 1, Kokchetavskii uiezd. Povtornoe izsledovanie 1907 goda, GUZiZ. Resettlement Administration, St. Petersburg.

Kuznetsov, V. K. (1910*a*), *Kirgizskoe khoziaistvo v Akmolinskoi oblasti.*, Vol. 5, Akmolinskii uiezd. Povtornoe izsledovanie 1909 goda, GUZiZ. Resettlement Administration, St. Petersburg.

Kuznetsov, V. K. (1910*b*), *Kirgizskoe khoziaistvo v Akmolinskoi oblasti.*, Vol. 4, Atbasarskii uiezd. Povtornoe izsledovanie 1909 goda, GUZiZ. Resettlement Administration, St. Petersburg.

Kuznetsov, V. K. (1910*c*), *Kirgizskoe khoziaistvo v Akmolinskoi oblasti.*, Vol. 2, Omskii uiezd. Povtornoe izsledovanie 1908 goda, GUZiZ. Resettlement Administration, St. Petersburg.

Kuznetsov, V. K. (1910*d*), *Kirgizskoe khoziaistvo v Akmolinskoi oblasti.*, Vol. 3, Petropavlovskii uiezd. Povtornoe izsledovanie 1908 goda, GUZiZ. Resettlement Administration, St. Petersburg.

LaFerrara, E. (2003), 'Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana', *American Economic Review* **93**(5), 1730–1751.

Leonardi, R., Nanetti, R. Y. & Putnam, R. D. (2001), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press, Princeton, NJ, USA.

Leunig, T., Minns, C. & Wallis, P. (2011), 'Networks in the Premodern Economy: the Market for London Apprenticeships, 1600–1749', *The Journal of Economic History* **71**(2), 413–443.

Ministry of Agriculture and State Property. Department of State Landed Property, (1900), *Ocherk rabot po zagotovleniiu pereselencheskikh uchastkov, 1893–1899*, Omsk.

Ministry of the Interior. Resettlement Administration, (1903), *Spisok pereselencheskikh i zapasnykh uchastkov, obrazovannykh s 1893 goda po 1-oe iiula 1903 goda v Stepnom kraie, Tobolskoi, Tomskoi, Yeniseiskoi i Irkutskoi guberniiakh*, number XIV *in* '[Spravochnyia] izdaniia pereselencheskago upravleniia', Staff of the Ministry of the Interior, St. Petersburg.

Molodykh, I. A. (1906), *Spisok pereselencheskikh i zapasnykh uchastkov, obrazovannykh po*

1-oe yanvaria 1904 goda v Akmolinskoi, Tobolskoi, Tomskoi, Yeniseiskoi guberniiakh (s kartami oznachennykh gubernii i oblasti), number XXII *in* '[Spravochnyia] izdaniia pereselencheskago upravleniia', Ministry of the Interior. Resettlement Administration, St. Petersburg.

Morrison, A. (2012), 'Sowing the Seed of National Strife in This Alien Region: The Pahlen Report and Pereselenie in Turkestan, 1908–1910', *Acta Slavica Iaponica* (31), 1–29.

Munshi, K. (2003), 'Networks in the Modern Economy: Mexican Migrants in the US Labor Market', *The Quarterly Journal of Economics* **118**(2), 549–599.

Munshi, K. & Rosenzweig, M. (2006), 'Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy', *American Economic Review* **96**(4), 1225–1252.

North, D. C. (1990), *Institutions, Institutional Change and Economic Performance*, Vol. 332 of *Political Economy of Institutions and Decision*, Cambridge University Press.

Nunn, N. (n.d.), 'The Long-Term Effects of Africa's Slave Trades'.

Platteau, J.-P. & Abraham, A. (1987), 'An Inquiry Into Quasi-Credit Contracts: The Role of Reciprocal Credit and Interlinked Deals in Small-Scale Fishing Communities', *The Journal of Development Studies* **23**(4), 461–490.

Roland, G. (2004), 'Understanding institutional change: Fast-moving and slow-moving institutions', *Studies in Comparative International Development* **38**, 109–131.

Rosenzweig, M. R. & Stark, O. (1989), 'Consumption smoothing, migration, and marriage: Evidence from rural india', *Journal of political Economy* **97**(4), 905–926.

Rumiantsev, P. P. (1910), *Kirgizskii narod v proshlom i nastoiashchem*, GUZiZ. Resettlement Administration, St. Petersburg.

Shakhmatov, V. F. (1964), *Kazakhskaia pastbishchno-kochevaia obshchina: Voprosy obrazovaniia, evolutsii i razlozheniia*, Izdanie Akademii Nauk Kazakhskoi SSR, Almaty.

Shcherbina, F. A. (1898), *Materialy po kirgizskomu zemlepol'zovaniiu, sobrannye i razrabotannye ekspeditsiei po issledovaniiu stepnykh oblastei.*, Vol. 1, Kokchetavskii uiezd, GUZiZ, Voronezh.

Shcherbina, F. A. (1900), *Ob'iasnitel'naia zapiska k dannym ob izlishnikh zemliakh, nakhodiashchikhsia v pol'zovanii kirgiz sievernoi chasti Kustanaiskago uiezda*, Orenburg.

Shcherbina, F. A. (1902*a*), *Materialy po kirgizskomu zemlepol'zovaniiu, sobrannye i razrabotannye ekspeditsiei po issledovaniiu stepnykh oblastei.*, Vol. 2, Atbasarskii uiezd, GUZiZ, Voronezh.

Shcherbina, F. A. (1902*b*), *Materialy po kirgizskomu zemlepol'zovaniiu, sobrannye i razrabotannye ekspeditsiei po issledovaniiu stepnykh oblastei.*, Vol. 11, Omskii uiezd, GUZiZ, Omsk.

Shcherbina, F. A. (1907), *Materialy po kirgizskomu zemlepol'zovaniiu, sobrannye i razrabotannye ekspeditsiei po issledovaniiu stepnykh oblastei.*, Vol. 3, Akmolinskii uiezd, GUZiZ, St. Petersburg.

Shcherbina, F. A. (1908), *Materialy po kirgizskomu zemlepol'zovaniiu, sobrannye i razrabotannye ekspeditsiei po issledovaniiu stepnykh oblastei.*, Vol. 12, Petropavlovskii uiezd, GUZiZ, St. Petersburg.

Smirnov, Y. A. (1899), *Otchet revizora zemleustroistva Smirnova po komandirovkie lietom 1899 g. v Akmolinskuiu i Semipalatinskuiu oblasti*, Ministry of Agriculture and State Property. Department of State Landed Property, St. Petersburg.

Smirnov, Y. A. (1900), *Dopolnenie k otchetu revizora zemleustroistva Smirnova po komandirovkie ego lietom 1899 g. v Akmolinskuiu i Semipalatinskuiu oblasti*, Ministry of Agriculture and State Property. Department of State Landed Property, St. Petersburg.

Special Commission on the Needs of Agriculture (1903), Vol. 57: Stepnoi Krai, *in* 'Trudy Mestnykh Komitetov o Nuzhdakh Sel'skokhozyaystvennoy Promyshlennosti', St. Petersburg.

Tatishchev, A. A. (2001), *Zemli i lyudi. V gushche pereselencheskogo dvizheniya (1906–1921)*, Russkiy Put, Moscow.

Tolybekov, S. E. (1971), *Kochevoe obshchestvo kazakhov v XVII–nachale XX veka. Politiko-ekonomicheskii analiz*, Nauka, Kazakh SSR, Almaty.

Townsend, R. M. (1994), 'Risk and Insurance in Village India', *Econometrica* pp. 539–591.

Udry, C. (1994), 'Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria', *The Review of Economic Studies* **61**(3), 495–526.

Vasilyev, D. V. (2018), *Bremya Imperii: Administrativnaya politika Rossii v Tsentralnoi Azii. Vtoraya polovina XIX veka.*, ROSSPEN, Moscow.

Wegge, S. A. (1998), 'Chain migration and information networks: Evidence from nineteenth-century hesse-cassel', *The Journal of Economic History* **58**(4), 957–986.

Wortman, R. S. (1976), *The Development of a Russian Legal Consciousness*, University of Chicago Press.

Zhilinskii, I., Vazhevskii, V., Kravtsev, I., Shiriaev, F., Kozyrev, A., Romanov, N., Przemycki, F., Tikhonovich, N. & Rakov, A. (1907), *Ocherk gidrotekhnicheskikh rabot v raionie Sibirskoi zhel. dor. po obvodneniiu pereselencheskikh uchastkov v Ishimskoi stepi i osusheniiu bolot v Barabie. 1895–1904.*, General Administration of Land Use and Agriculture. Division of Land Improvement, St. Petersburg.