CenEA Working Paper Series

WP03/15

How History Matters for Student Performance. Lessons from the Partitions of Poland

Paweł Bukowski

Abstract:

This paper examines the effect on current student performance of the 19th century Partitions of Poland among Austria, Prussia and Russia. Despite the modern similarities of the three regions, using a regression discontinuity design I show that student test scores are 0.6 standard deviation higher on the Austrian side of the former Austrian-Russian border. This magnitude is comparable to the black vs. white test score gap in the US. On the other hand, I do not find evidence for differences on the Prussian-Russian border. Using a theoretical model and indirect evidence I argue that the Partitions have persisted through their impact on social norms toward local schools. Nevertheless, the persistent effect of Austria is puzzling given the historical similarities of the Austrian and Prussian educational systems. I argue that the differential legacy of Austria and Prussia originates from the Austrian Empire's policy to promote Polish identity in schools and the Prussian Empire's efforts to Germanize the Poles through education.

Key words: institutional persistence, identity, social norms, quality of education, the Partitions of Poland

JEL classification: N30, I20, O15, J24

Centre for Economic Analysis
CenEA
www.cenea.org.pl

How History Matters for Student Performance. Lessons from the Partitions of Poland *

Paweł Bukowski †

This paper examines the effect on current student performance of the 19th century Partitions of Poland among Austria, Prussia and Russia. Despite the modern similarities of the three regions, using a regression discontinuity design I show that student test scores are 0.6 standard deviation higher on the Austrian side of the former Austrian-Russian border. This magnitude is comparable to the black vs. white test score gap in the US. On the other hand, I do not find evidence for differences on the Prussian-Russian border. Using a theoretical model and indirect evidence I argue that the Partitions have persisted through their impact on social norms toward local schools. Nevertheless, the persistent effect of Austria is puzzling given the historical similarities of the Austrian and Prussian educational systems. I argue that the differential legacy of Austria and Prussia originates from the Austrian Empire's policy to promote Polish identity in schools and the Prussian Empire's efforts to Germanize the Poles through education.

JEL Classification: N30, I20, O15, J24

*I thank Sascha O. Becker, Volha Charnysh, Gregory Clark, Tomas Cvrcek, John S. Earle, Irena Grosfeld, Hedvig Horvát, Gábor Kézdi, Jacek Kochanowicz, Attila Lindner, Christina Romer, Ruth M. Schüler, Tamás Vonyó, Jacob Weisdorf, Agnieszka Wysokińska, Noam Yuchtman, the participants of seminars at Central European University, University of California at Berkeley, University of California at Davis, Warsaw School of Economics, Ifo Center for the Economics of Education, CenEA and FRESH workshops in Warsaw and Canterbury, WEast workshop in Belgrade, European Historical Economics Society Summer School in Berlin for their comments and suggestions. I acknowledge the hospitality of the Department of Economics, CLE, IRLE, and BEHL at UC Berkeley. I gratefully acknowledge financial support from the Review of Economic Studies studentship. All errors are mine. The previous version of this paper was publicized through the Berkeley Economic History Lab Working Papers series.

†Central European University, The Department of Economics, Nador u.9, 1051 Budapest, Hungary. E-mail: bukowski_pawel@phd.ceu.edu

1 Introduction

An intriguing idea in recent economic and historical research is that modern economies are affected by past institutions even after the institutions have ceased to exist (Acemoglu and Robinson, 2008). In the case of education, the historical investments in public goods and property rights institutions have been shown to affect current educational attainment, provision of schools and literacy levels (Banerjee and Iyer, 2005; Huillery, 2009; Iyer, 2010). We know less about the underlying mechanisms and whether they depend on social context. It has been argued, for instance, that universal schooling might level the historical differences in educational outcomes (Dell, 2010). In this paper, I show that two historical parts of Poland, which had similar past educational system and provision of public education, had - relative to a control region - very different long run effects on current student performance. I show evidence highlighting the role of social norms toward local schools as a key channel of persistence (Akerlof and Kranton, 2010; Sakalli, 2014). I also argue that the interaction between national identity and institutions created different social norms toward local schools in the two historical parts, generating the difference in student performance today.

Specifically, I analyze the Partitions of Poland (1815-1918) among Austria, Prussia, and Russia (see Figure 1), as a laboratory to investigate how history matters for student performance. The comparison of geographic characteristics and the historical literature suggest that the former borders between the Empires were not drawn to reflect any pre-existing socio-economic, historical, geographic or ethnic divisions (Wandycz, 1974; Becker, Boeckh, Hainz and Woessmann, 2014b; Grosfeld and Zhuravskaya, 2015). Consequently, I will argue the Partitions of Poland provide an exogenous variation in institutional heritage in modern Poland. The three partitions differed significantly. However, in terms of educational systems, the Austrian and Prussian institutions were very similar as the former was copied from the latter (Cohen, 1996; Lamberti, 1989). The Austrian and Prussian system was financed from local taxes, had compulsory elementary and optional secondary education and shared similar curricula and pedagogical methods. As a result, provision of public education was comparable in the Austrian and Prussian partitions. The Russian educational system, in turn, practically did not exist in the 19th century (Snyder, 2006).[1] The three regions of interest are now within Poland, are ethnically homogeneous and have the same modern educational and legal systems.

Using regression discontinuity design I compare test-measured performance of students in municipalities at the two sides of the former border between Austria and Russia. I show that student test scores on the Austrian side are 0.61 standard deviation higher, which is similar to the black vs. white achievement gap in the US. On the other hand, I do not find evidence for differences on the Prussian-Russian border. These results provide evidence history matters in the long-run and are consistent with other studies documenting long lasting effects of historical heritage (e.g. Acemoglu, Johnson and Robinson, 2001; Basten and Betz, 2013; Dell, Lane and Querubin, 2015).

[1] The Russian system had no compulsory elementary schooling, no coherent organization of a school network and no political will for expanding education.

There are many potential channels through which the Partitions has affected the current student performance, but the data availability does not allow to identify all of them. I highlight that people living in the former Austrian Empire have inherited positive social norms toward local schools, which lead to a higher schooling effort and thus increase the performance of students. The social norm channel has been underlined in general studies (Karaja, 2013; Becker et al., 2014b; Grosfeld and Zhuravskaya, 2015) and in the context of educational outcomes (Sakalli, 2014; Feir, 2015). I provide three pieces of evidence to support it. Firstly, I show that the effect of the Austrian Empire is larger on the low stake exam score than on the high stake, which is consistent with a social norm-based model of student effort (Akerlof and Kranton, 2002). Intuitively, social norms toward local schools matter more for the low stake exam because there is no universal motivation to obtain a high score. Secondly, I use survey data on proxies for social norms to show that people from the former Austrian Empire are more likely: to choose education as first or second priority in governmental spending; to say that education is crucial for a decent life and to select family tradition as an important determinant of school choice. Finally, I show that the Austrian partition has a positive and large effect on kindergarten attendance that cannot be explained by the historical supply of kindergartens. At the same time, I show that other channels, in particular skill-biased migrations, are unlikely to explain my results. The historical migration patterns do not suggest any strong selection in and out-migration to/from the Austrian partition. To evaluate the present-day migration I adjust the modern data to match a hypothetical extreme skill-biased migration case and check whether I still document a sizable effect of the Austrian partition. This exercise shows that the current migration is not the major force responsible for the observed effect.

Why social norms differ in the Austrian and Prussian partitions is puzzling given given that the former was not economically superior over the later and both Empires had almost identical educational systems and similar provision of public education.[2] I argue that the differential legacy of Austria and Prussia originates from the different interaction between educational institutions and Polish identity. While the Prussian state used these institutions mainly to Germanize Poles (e.g. through the German language of instruction), the Austrian state used them to support Polish identity (e.g. through the Polish language of instruction) (Cohen, 1996; Lamberti, 1989).[3] Because of the historical attitude of the Polish population toward the educational systems, positive social norms toward education may have been more likely to emerge in the Austrian partition. These could be then transmitted through generations and still affect student and parental effort. Consistent with this hypothesis, Steele and Aronson (1995) and Akerlof and Kranton (2010) provide theoretical and empirical evidence that identity is associated with social norms affecting an individual's schooling choices, school-student relationships and student achievements.

I provide suggestive evidence for the importance of interaction between institutions

[2]Consistently, Michalopoulos and Papaioannou (2013) and Grosfeld and Zhuravskaya (2015) show that not all institutions influence social norms and matter in the long run.

[3]Russia also used education as a tool to Russify the population and so the language of instruction was Russian (Snyder, 2006).

and identity. Using the historical data on the 19th century educational outcomes in Austria and Prussia, I correlate the historical elementary school enrollment with the current student performance. The results show that the correlation is null in the areas which are in the former Austrian partition, but strongly negative in the former Prussian partition. These estimates are robust to the inclusion of geographical and socio-economic covariates, yet they might be not causal. However, assuming that the remaining bias is the same in both regions, the historical expansion of the education system has more positive effect on the current student performance in the former Austria than in the former Prussia. This is in line with the proposed hypothesis, as in the Austrian Empire there was a *positive* interaction between identity and institutions. Hence, the social norms affecting student performance may have been more likely to emerge in municipalities with a larger attachment to the historical Austrian educational system. Alternatively, because of the *negative* interaction between the institutional quality and identity, the more intensive historical exposure to the Prussian education might lead to a stronger opposing social norm toward the educational system. This norm leads to a lower schooling effort and thus decreases the performance of students.

Overall, contributions of this study are threefold. Firstly, I show that history matters for student performance and it accounts for a sizable gap in educational achievements. Secondly, I provide evidence that history has persisted through its impact on social norms toward local schools. Finally, I propose a source of persistence based on the interaction between institutions and identity.

The studies that are closest to mine are Grosfeld and Zhuravskaya (2015),Wysokinska (2011) and Becker et al. (2014b).[4] Grosfeld and Zhuravskaya (2015) find a persistent effect of the Partitions of Poland on the level of religiosity, belief in democratic values and rail-road infrastructure, but not on income, industrial production, the share of people with secondary education, corruption and trust in government institutions. Consistently with my study, the authors argue that the inter-generational transmission of social norms can shape political and religious preferences, even though the majority of differences between the partitions have been smoothed out by economic factors. This study shows that a historical institution affects behavior differently in different domains. In contrast, I show that it affects behavior in the same domain differently in different places. Wysokinska (2011) provides a general impact of the Prussian Empire and finds a positive effect of the German administration on general trust, income and turnout for referenda. Finally, Becker et al. (2014b) point out that, among the Central-Eastern European countries, the Habsburg Empire is associated positively with trust toward the local state and negatively with acceptance of corruption. All the mentioned studies use regression discontinuity designs in their identification strategies.

In addition to these, my results are partially consistent with Herbst (2004) and Herbst and Rivkin (2012), who analyze determinants of the distribution of the exam scores in Poland. In particular, they regress the exam scores for all municipalities in Poland on a set of modern-day control variables and the partitions dummies. They find that relative

[4]For the literature in Polish see: Hryniewicz (2003), Chuminski (2008).

to Warsaw, the dummy for the former Austrian part has the largest magnitude, and for the Prussian Empire the lowest. However, they do not use regression discontinuity design and the current covariates are likely to be endogenous resulting in biased estimates (Angrist and Pischke, 2008). The authors also do not empirically identify the channels of persistence of the Partitions of Poland.

The paper is organized as follows. In Section 2, I present the historical overview of the Partitions of Poland and look in more detail at the educational system in each Empire. In Section 3, I describe the data, research methodology and show the effect of Partitions of Poland on the performance of students. Section 4 identifies the channels of persistence. Section 5 discusses the sources of persistence. Finally, Section 6 concludes and discusses policy implications.

2 Historical Overview

This section describes in more detail the Partitions of Poland and situation of the Poles in the 19th century educational systems in Prussia, Austria and Russia. For readers not interested in historical details, it is sufficient to read the summary at the end of this section. The summary also contains information about the modern educational system in Poland.

The Partitions of Poland took place in three parts, during the second half of the 18th century and put an end to a two-hundred year old Polish-Lithuanian Commonwealth.[5] Due to the Partitions, Poland was removed from the map of Europe for 123 years and came back into existence after World War I. The first annexation of the Polish lands by the Russian Empire, the Kingdom of Prussia and the Habsburg Austria took place in 1772 and as a result Poland lost almost one-third of its territory and 4.5 million inhabitants. In 1793 Russia and Prussia conducted the second partition, which further decreased the territory and finally in 1795 all three Empires absorbed the rest of the remaining country. Thanks to Napoleon I this situation did not last for long. In 1807 he conquered the Central and Eastern parts of Europe and established the Duchy of Warsaw - a Polish state controlled by one of Napoleon's allies. However, the Duchy survived only seven years as the defeat of Napoleon I in 1814 brought back the situation before the Napoleonic Wars.

The new border between the partitions was established during the Congress of Vienna in 1815 after which they remained generally unchanged until the end of World War I. During the first decades after the Congress, the Russian and Prussian administrations were not systematically oppressive toward the Poles. The Congress Kingdom and the Grand Duchy of Poznań - newly created states controlled by Russia and Prussia respectively - experienced some level of freedom and gave the Poles hope that independence was within their reach. In the Congress Kingdom this had lasted until the unsuccessful

[5]For a more detailed historical description of the Partitions of Poland and debate about the sources of the Commonwealth failure, see Davies (2005a), Davies (2005b).

uprising against Russia in 1830, after that the Poles were repressed and Russified[6] until the end of World War I. In Prussia, the situation of the Poles worsened in the 1870s when Otto von Bismarck introduced *kulturkampf*.[7] Differently from the other partitions, the Poles under Austrian occupation had relatively less freedom during the first part of the 19th century, but it changed after 1867 when the Austrian administration took a more tolerant and multicultural approach in their policy. Language freedom was one of the most significant expressions of this. Polish was the official language of the Galician administration (Galicia is part of Poland and Ukraine, which was under the Habsburg rule) and could be used as the language of instruction in schools. Contrary to this, in the Russian and Prussian parts, from the second part of the 19th century the usage of Polish was limited both in administration and education.

In terms of the socio-political situation, the Prussian and Austrian partitions were more favorable to the self-organization of the Poles. Both Empires introduced on the Polish lands a bureaucratic system with a strong "administrative ethos" (Gillis, 1971; Becker et al., 2014b). The Prussian state was above all a state of law and even though the administration was discriminating the Poles, the people created institutions such as agricultural societies, credit institutions, reading rooms, newspapers and educational circles to support Polish economic activity and defend the national identity. Ethnic tolerance and freedom in the Austrian part resulted in numerous associations, newspapers and institutions spreading and preserving Polish culture. Two universities in Galicia, the Jagiellonian University in Cracow and Lviv University, played a very important role in the development of Polish intellectual life. They also attracted Poles from the other partitions[8] and by this contributed to the preservation of the nation's intellectual heritage. All these were in contrast with the situation of Poles under Russian rule, where bureaucracy was inefficient (Burke, 1979) and most forms of self-organization were forbidden and fought by the Tsarist administration.

The best economic situation was in the Prussian zone. The authorities carried out many reforms there. The most important of these was the manumission, that is the peasants could become owners of the land after repaying the nobility. Rising demand for agricultural products induced changes in agricultural technology (crop rotation), fertilizers were applied, and the wealthier farmers were buying machinery. As a result, the agriculture, rather than industry, was the main drive of the economic progress in the Grand Duchy of Poznań. Economies of the other partitions were different. In the Russian zone it was industry that developed the most. The clusters of textile industry were created in Łódź and Białystok. Warsaw became a modern city with its sewers, streets, gas lighting, and power plant switchboard. Economic progress, however, did not improve the well-being of workers who had to work long hours (14 hours) for low wages

[6]The most important expressions of Russification were ban on using the Polish language in public spaces, forbidding teaching of the Polish language and the history of Poland, promotion of the Russian Orthodox faith combined with repressions toward the Catholic Church. Additionally, the tsarist government deported many students and intellectuals involved in secret polish societies and fraternities (Wandycz, 1974; Snyder, 2006).

[7]A policy direction, which consisted of measures against the Catholic church and the Polish nation.

[8]This migration was small and limited to Cracow and Lviv. I discuss it in Secion 4.2.

and in unsafe conditions. The delayed manumission reforms, which were introduced only during the second half of the 19th century, contributed to the relative backwardness of the agriculture in the Congress Kingdom. Nevertheless, the worst economic situation was in the Austrian part. Before the end of the 19th century Galicia had not been industrialized and the agriculture was under invested and parceled. Consequently, people had experienced one of the worst poverty rates in all of the Habsburg Empire, and at the beginning of the 20th century over two million Galicians emigrated abroad to escape the bad economic conditions.

In the following subsections I discuss the situation of the Polish minority in Prussia, Austria and Russia, in the context of the 19th education systems.

2.1 The Prussian Education System

In 1763 the Prussian state created an education system which became a model for numerous other countries, including the US, Japan and Austria. Although it was changed many times during the 18th and 19th centuries, the core of the system was the obligatory elementary school (*Volksschule*) followed by various types of secondary school. Despite its centralized design, the financing of the education was based on local taxes and municipal school boards managed the school operation (Cinnirella and Schueler, 2015). Wilhelm von Humboldt, who in 1809 was appointed the Prussian Minister of Education, developed the idea of universal and compulsory education. Thanks to him, the schooling system became perceived not only as a source of specialists, but also gained an universal aim of the general intellectual development of society.

Beside its modernity and universal character, until 1870 the elementary school was practically a domain of the church (both Protestant and Catholic). Most schools were confessional, and religion was the main subject in the Prussian curricula. The reformative movements of 1848 were trying to emancipate the school from the church influence, but not much was changed. On the one hand, the state was trying to promote a secular and nation-oriented[9] education. On the other, it was afraid that taking too much power from the church would motivate it to create a competitive network of private schools. Two decades later, Adalbert Falk - the Minister of Ecclesiastical Affairs during *kulturkampf* - implemented a new set of secularization reforms. They included the limitation of the church's influence, professionalization and secularization of the school inspectorate. Yet the impact of the reforms was limited, as the clergy retained its strong position.

Nevertheless, the policy turned out to be very important in the Polish context. The local Catholic Church[10] helped to cultivate the Polish national identity more than any other secular movement. Consequently, *kulturkampf* was done more consistently on the Polish lands than anywhere else (Lamberti, 1989). In addition to this, from 1870 the Prussian state executed repressions on a much larger scale than it had done before, in

[9]Understood as the German nation.

[10]The Protestant church was also affected by *kulturkampf*, but because of its special role in the Prussian state, to a much smaller extent than the Catholic.

particular, it banned the use of the Polish language in administration and education, forbidden cultivation of the Polish traditions, discriminated Polish workers on the labor market and deported Poles and Jews, who did not have the Prussian citizenship.

The most important change for education was the language of instruction. In 1822 the Prussian state permitted the use of the Polish language of instruction in the regions with Polish population. This lasted until 1870, when *kulturkampf* redefined the role of elementary education. As Marjorie Lamberti states: "Prussian state officials looked to the *Volkschule* to serve as an instrument of Germanization. The school's function was not to only teach Polish children to speak German but also acculturate them into the German nation" (1989, p.109). As a result, in 1873 German was introduced in the Grand Duchy of Poznań and Eastern and Western Prussia as the language of instruction starting with the first two years of schooling. At the same time, Polish was permitted only during the religion classes and final exams.[11] When in 1901 Polish was banned completely, students and parents of Września started protest. Soon it turned into a massive strike, which included around 75 thousand students from 800 schools. Even though the scale of protests surprised the Prussian government and some politicians were calling for the revision of the anti-Polish policy, Heinrich Konrad von Studt - the Minister of Education - retained the policy. "This policy bred germanophobia and a repugnance for the school in Polish families" (Lamberti, 1989, p.109). But the language of instruction was not the only reason why Polish parents opposed the educational system.

The educational inequalities and feelings of unfairness were further reasons. The introduction of German as the language of instruction implied that the teachers had to teach in a language in which they did not always have the required proficiency. Moreover, the students from Polish speaking families had to first learn German, which meant less time for the other classes. Finally, the Polish schools were systematically under-financed compared to the German ones (Cinnirella and Schueler, 2015). The average student teacher ratio on the lands with the Polish population was 93:1, while in the rest of Prussia it was 60:1 (Lamberti, 1989, p.129). The situation was especially visible in Poznań, where a disproportional share =of the public money went to the German schools. All these translated into a lower quality of the Polish schools[12] and raised feelings of unfairness among Polish parents.

The situation of teachers was also ambiguous. As pointed out by Lamberti (1989), during the Schools Strike, the Polish teachers were generally not willing to support the parents' demands. They stood on the Prussian administration side because they were afraid of losing their jobs. This in turn led to acts of hostility toward the teachers: "[t]he Polish press rebuked the teacher for currying favor with the school inspectors and promoting the use of German in order to obtain bonuses. In public places the teachers were insulted, threatened and assaulted" (Lamberti, 1989, p.146). The parents not only distrusted and fell in conflict with the institution of elementary school but also with its personnel.

Finally, on the Polish lands the Prussian government was more active in introduc-

[11] However local governors could order exclusive teaching in German.

[12] Still it was much better than in Russia or Austria, see for example illiteracy rates in Figure 2.

ing educational reforms aimed against the church influence. The most profound were introduction of the interconfessional schools[13] and the secularization of the school inspectorate. From the very beginning, the Polish population viewed the innovation with distrust. As Lamberti claims: "The interconfessional school policy further alienated the Polish people from the school administration. [...] (they) had good reasons to believe that the interconfessional schools were being opened for the purpose of Germanizing the Polish youth" (1989, p.115).

The German language of instruction, inequalities, the role of teachers and the interconfessional education motivated the hostility toward the education system among the Polish families living in the Prussian Empire. Yet, in comparison to the other parts of Poland, the system was effective. Law enforcement was widespread and most of the children who attended the elementary school were taught how to read and write. This was partially because treating education as a tool of Germanization additionally motivated the administration to execute the compulsory schooling. As such, the Prussian educational system combined effective institutions with the set of anti-Polish regulations.

2.2 The Austrian Education System

In his comprehensive analysis of the 19th century education in Austria, Gary B. Cohen (1996) emphasizes that the institutional design of the Austrian education system was to a large extent a copy of the Prussian model. Already in 1781 Joseph II established the principle of mandatory primary education, however until 1848 the education system mainly served as a training field for administration officials (the Emperor Francis I used to say: "I need no learned men; I need only good officials"). The People's Spring movement brought the Humboldtian model of education and in 1850 Leo Thun, the Minister of Education, initiated a period of intensive reforms, which greatly modernized the education system. The strongest adherent and executor of the reforms was the faction of German Liberals in the Austrian Parliament, who patterned their ideas on the Prussian model. Although delayed by few decades, the amendments were paralleling the developments in Germany. The idea of local-tax funded elementary school (equivalent of Prussian *Volksschule*), which was obligatory until the age of 14, was fully introduced following the 1867 reform and the General Primary School Law of 1869. Also the secondary and higher education were modeled on the Prussian system (including the curricula), as Cohen states:

> The Austrian reformers of the late 1840s and 1850s adopted much of the early nineteenth century German model of academic secondary and higher education. [...] During the late nineteenth century, the discourse of the Austrian government officials and educators on such matters was much the same as that of their counterparts in Germany. The Austrians identified many of the same problems regarding curricula and the rapid growth in secondary and higher education as did their German counterparts (1996, p.259-260).

[13]Interconfessional schools (also called mixed) gather students from different religious groups.

Nevertheless, the systems differed in one important aspect. While in Prussia education was the main tool of Germanization, in Austria it was seen as a tool to promote national identities. However, it was not like this from the beginning. During the first part of the 19th century, the official language of instruction at all stages of education was German. Only in 1850 did the reformative movement introduce Polish at the primary education level. Still, as reported by Cvrcek and Zajicek (2013), in 1865 the local elites favored public education only if it was in German. It changed after 1867 when the Austrian administration took a multicultural approach in their internal policy. The second wave of reforms carried out by the German Liberals extended the Polish language of instruction to secondary and higher education. Thanks to this, three universities in Galicia played an important role in the preservation of nation's heritage and development of Polish intellectual life.

Another important aspect of Austrian education was its inclusiveness. The expansion of the elementary and secondary school network was possible thanks to the proactive attitude of local governments and voluntary associations. The growing demand for education of previously uneducated groups resulted in a numerous grass-root educational initiatives. Non-German speaking ethnic groups and the Jewish people had greater aspirations toward education than the Germans. Also new lower middle classes, for instance children of independent business owners, were considerably more attracted by the possibilities offered by education than the old elite. This was especially visible in the Polish part of the Austrian Empire, where the agriculture was backward and extensively parceled. The beginning of the 20th century saw a rapid growth in elementary and secondary education in Galicia, the share of elementary students in population almost tripled between 1880 and 1910 (GUS, 2003) (see Table 2), whereas the secondary enrollment ration increased by 120% (in the German speaking lands it increased by 52%) (Cohen, 1996).[14] As pointed out by Cohen (1996, p.257) "[b]y 1910 the Polish speaking share of Austrian enrollments significantly exceed the Polish speaking share of the Austrian population". There was also a strong popular and political pressure to open advanced education to children from poorer strata. At the same time, however, the literacy levels and school's attainment was still lower there than in the Prussian Partition or other parts of the Austrian Empire.[15]

The class instruction in Polish, broadening the access to education and poverty caused that Poles living in Galicia saw education as the main means for preserving their national identity and improving their material conditions. Even though the law enforcement and quality of institutions were not as good as in the Prussian Empire, the system managed to create positive relations with the citizens.

[14]The secondary enrollments analyzed per thousand inhabitants in the Polish speaking lands of the Habsburg Empire : 1880 - 2.74, 1890 - 2.78, 1900 - 3.77, 1910 - 6.05; German speaking lands: 1880 - 3.88, 1890 - 4.04, 1900 - 4.61, 1910 - 5.88 (Cohen, 1996, p.141)

[15]Cohen claims: "In the 1870s and 1880s the majority of of school aged children in Galicia [...]did not attend Volksschulen. In 1880 only 21% of the population 6 years or older could read in Galicia. In 1910 83.5% of over 11 years old population of Austria was literate while in Galicia this number was 58%." (Cohen, 1996, p.64)

2.3 The Russian Education System

The Tsarist administration followed the path of educational development initiated by Peter I and Catherine II almost until the end of the 19th century. Beside high investments in universities and growing numbers of enrolled students in elementary schools, the ruling class did not accept the Humboldtian approach to education. Sergei Uvarov, the Minister of Education (1831-1849) during the rules of Nikolai I, may be the best example. He laid the foundations for the modern and high quality higher education in Russia[16] but clearly opposed broadening and developing education for people from lower strata. He "believed that excessive education would only create dissatisfaction among the peasantry" and "the lower classes had to be protected from too much knowledge." (Kassof, 2004). This approach was also visible in other aspects of life in the Russian Empire and might have been partially responsible for the dissatisfaction of people, which led to the Bolsheviks Revolution in 1917.

The other problem was the chaotic organization of the school system. There was no obligatory schooling and the educational policy was inconsistent, as the Ministry of Education did not control the network of schools.[17] Lacking the central design and organization, the system was characterized by class-based duality, with separate curricula for students from upper and lower stratum. Consequently, the illiteracy levels were very high: in 1917, only 70% and 30% of urban and rural population respectively could read and write.[18]

The situation was especially bad on the Polish lands (the Congress Kingdom). The lack of educational institutions was accompanied by very intensive Russification and the repression of the Poles[19] (Chubarov, 2000). Polish society under the Tsarist rule not only was underdeveloped in terms of education but also had to fight for its national identity. For instance, due to the repression, which took place after the November Uprising in 1830, the number of secondary school students was reduced by 50% until 1855 (Snyder, 2006).

Many studies underline the rapid development of education in the Tsarist Russia, especially at the end of the 19th century. This becomes undoubtedly true once we think about the general situation of the Russian society during, for example, the Napoleonic Wars. Nevertheless, the Congress Kingdom was one of the most advanced parts of the Russian Empire in terms of economic and social development. Once compared with the other parts of Poland, one may argue that its educational potential was wasted to a large extent.

[16] On the other hand, he is responsible for the closure of the University of Vilnus after the November Uprising in 1830 (Whittaker, 1984).

[17] (Kassof, 2004) estimates that "sixty-seven different types of primary schools [existed] in Russia in 1914"

[18] As pointed out by Bowen (1962, p.23), during World War I, "literacy was so rare that most Russian troops were unable to write home, even if their families could read".

[19] Interestingly, the policy of the Russian Empire toward other nations was not always that harsh. Alexander II hated in Poland, has a monument in Helsinki.

2.4 Summary

Table 1 summarizes the main differences between the partitions. Developed agriculture, modern bureaucracy and strong law enforcement characterized the Prussian partition. The later allowed self-organization of the Poles, which contributed to the preservation of the Polish culture, threaten by the Prussian's attempts to Germanize the Poles. In the Russian partition industrialization led to modernization and development of cities. But the weak law enforcement and the anti-Polish orientation of the Tsarist policy undermined position of the Poles. In the Austrian partition, backward agriculture and industry were responsible for harsh socio-economic conditions. However, the Austrian administration developed an effective bureaucracy apparatus and since 1860's this was the only partition with a significant autonomy given to the Polish population.

In terms of educational systems, the Austrian and Prussian institutions were very similar as the former was copied from the latter. The Austrian and Prussian system was financed from local taxes, had compulsory elementary and optional secondary education and shared similar curricula and pedagogical methods. However, while the Prussian state used these institutions mainly to Germanize Poles (e.g. through the German language of instruction), the Austrian state used them to support Polish identity (e.g. through the Polish language of instruction). Consequently, the Poles under Prussian rule opposed the educational system and were hostile toward the school personnel (especially teachers). Remarkably, massive school strikes were organized by Polish parents, the largest one took place in 1901 when 70 thousands Polish students refused to go to school. The Russian educational system, in turn, practically did not exist in the 19th century.

The differences in the educational outcomes between the three partitions are documented in Table 2. School enrollment in the Prussian part in 1864 was as high as 93%, while in the Austrian part it was significantly lower throughout the 19th century, but quickly converged to the Prussian level by 1914. Notably, at the outset of WWI, in the Russian part less than 25% of the school age population attended a school. Similarly, the provision of public schools in the 1910's was practically the same in the Austrian and Prussian partitions, in the former on average there was one school per $13km^2$, in the later one school per $10km^2$. Contrary to this, in the Russian partition there was one school per $27km^2$. As a result, after Poland gained independence in 1918, on the formerly Russian lands the illiterate population was as high as 65 percent, whereas in the former Prussia it was less than one percent. The illiteracy levels in 1931 are depicted in Figure 2. Regions in the West had the lowest level of illiteracy, moderately higher in the South (except for the presently Ukrainian parts) and highest in the Central and Eastern parts of Poland. These differences were to a large extent smoothed after World War II when the 8-year education became obligatory in all of Poland (Meissner and Majorek, 2000). Yet social norms toward education, affecting student performance could not be easily smoothed.

Today, the Polish comprehensive and compulsory education system consists of 6 years of elementary school, which is then followed by 3 years of *gimnazjum*. The admission to the comprehensive schools is based on catchment areas, which means that every student living within this area has a right to attend a given public school. However, parents

may request an alternative school, but its principal has a right to reject the application. During the comprehensive education, students are examined by the two standardized, externally graded and obligatory examinations: a low stake after elementary school (6th grade) and a high stake after *gimnazjum* (9th grade). The later serves as a basis for the admission into the higher secondary education, which is a first part of tracking. Students can choose a track (academic, mixed or vocational) and apply to any high schools, but the admission is not granted.

3 The Partitions of Poland and Student Performance

3.1 Data

My analysis draws on comprehensive municipality-level registry data on obligatory examination scores for the period 2005 - 2011, published by the Central Examination Board of Poland. The available exam scores are from a low stake general 6th grade exam (taking place after elementary school) and a high stake mathematics-science 9th grade exam (after lower secondary school). While the former serves mainly an informational purpose, the later matters in the high school admission process and thus motivates students (and their parents) to obtain the best score. In addition to this, a set of socio-economic control variables are available at the municipality level from the Central Statistical Office of Poland and the System of Educational Information. Geographical and climate data come from the *WorldClim.org* project (Hijmans, Cameron, Parra, Jones, Jarvis et al., 2005). For the full description of the available variables see Table A1.

Descriptive statistics for *rural* municipalities located at most 50km from the borders are presented in Table 3. These variables are for the present period, and, as such, are endogenous with respect to the Partitions of Poland.[20] The border areas of the former Russian Partition seem to have the worst socio-economic situation, as the rate of unemployment is the highest, the expenditures are the lowest and the migration balance is negative. The municipalities which were under the Prussian rule are characterized by high share of employment in agriculture [21], high share of people aged 0-18, positive migration balance and low level of unemployment. The situation in the former Austrian zone is similar to the former Prussian, except a low importance of agriculture and high population number and density.

The rural borderlands of the former Austrian partition has the best educational outcomes (except the number of additional classes and the level of scholarization), even though these lands are not necessary better in the case of the other socio-economic characteristics. Importantly, it has also higher educational spending per capita, but this difference disappears when the general spatial trends are accounted for (see Table 6). The former Prussian and Russian borderlands have similar level of achievements, but

[20] In other words, they might reflect the effect of the Partitions of Poland

[21] The agriculture practice on the former Prussian lands is the most efficient in Poland. It is based on large, business-oriented farms, which are not very common in the rest of the country.

the former have the largest classes and highest number of additional lessons. The high performance of students from the former Austrian partition is also confirmed by Figures 3 and 4, which show the spatial distribution of the 2011 6th and 9th grade exams for the whole sample of municipalities. It can be clearly seen that the territory of the former Austrian Empire is a cluster of high-performing municipalities.

3.2 Empirical Strategy

Straightforward comparison of schools in the former Prussian, Austrian and Russian partitions neglects other differences between these areas, which are largely unobserved, and may lead to biased estimates of the effect of the Partitions of Poland. It is possible, for example, that proximity to Germany or Ukraine matters (through e.g. trade possibilities and resulting returns to education) and the further we go south-east, the exam scores are increasing and we mistakenly conclude that this is due to Austrian rule. To solve this problem, I follow Dell (2010) and employ a geographical two-dimensional regression discontinuity design, which evaluates the effect of the Partitions of Poland by focusing on a discontinuous jump at the borders.[22] To control for the potential confounding effects of a geographical location, I narrow the analysis only to areas located close to the partitions borders and include into a regression a polynomial of latitude and longitude. The model can be written as:

$$y_{it} = \alpha + f(location_i) + \beta D_i + \gamma G_i + \epsilon_{it} \tag{1}$$

where i indexes municipality and t indexes year. $f(location_i)$ is a polynomial of latitude and longitude, the dummy D takes value 1 for the former Russian areas and value 0 for either the Austrian or Prussian, G_i are time-invariant geographical controls (altitude, precipitation and temperature), and ϵ_{it} denotes idiosyncratic shocks. The two outcome variables are the standardized (Z-score) 6th grade exam score and the standardized mathematics and science 9th grade exam scores.[23] They are available from 2005 to 2011. The sample consists of municipalities, which are located within a given distance to the borders (the bandwidth).[24] I pool the data and estimate the model using the Random Effect estimator, and, as a robustness check, I also estimate it year-by-year using OLS.

The regression discontinuity framework requires a proper specification of the polynomial $f(location_i)$ and the bandwidth. There are no theoretical arguments for any specific order, therefore I report results for linear, quadratic, cubic and quartile polynomials. Nevertheless, the Akaike information criteria (Lee and Lemieux, 2010) favor the quadratic polynomial, so I consider it as a baseline model. The bandwidth selection is based on the trade-off between the sample size and internal validity. For my baseline

[22] For more about the geographical regression discontinuity design see Keele and Titiunik (2011), for general discussion about the regression discontinuity framework see Imbens and Lemieux (2008), Lee and Lemieux (2010), Angrist and Pischke (2008).

[23] The variables are standardized (demeaned and divided by standard deviation) for each year separately.

[24] The author used the GIS data to calculate the distance between the municipality centroid and the corresponding border.

specification I choose 50km bandwidth and in the robustness section I also report results for municipalities located at most 75km and 100km.

3.3 The Borders under Investigation

The key assumption for the regression discontinuity design to provide the causal effect of the Partitions of Poland is that exogenous variables, influencing educational performance are smooth at the border. There is a consensus among historians that that the borders of interest[25] were not drawn to reflect the pre-existing socio-economic, historical, geographic or ethnic divisions (Wandycz, 1974, p.11). Nevertheless, I exclude Silesia and Eastern Prussia from my analysis, because during the interwar period (1918-1945) they belonged to Germany and were a destination point for the massive post WWII resettlement of Poles from the territories of modern Belarus, Lithuania and Ukraine.[26] Without these regions, I ensure that the observed difference between the areas of interest is due to the Partitions of Poland, not some later historical event. As a result, my sample consists of areas which had similar history *before* and *after* the Partitions, were ethnically homogeneous and are now within the territory of Poland. Figure 1 depicts the partitions borders layered on the modern boundaries of Poland, the solid line representing the borders under investigation, the dashed line marking the excluded parts. Please note that by excluding Silesia I cannot directly compare the borderlands between the Austrian and Prussian partitions. Figure 5 presents the total area under investigation, namely the rural and urban municipalities located at most 50km from the borders of interest.

In my baseline specifications I focus on the rural areas because of two reasons. Firstly, the current migration of people from the rural to the urban areas, which ignores the Partitions borders, blurs interpretation of the Partitions of Poland effect in the urban areas. Secondly, large cities (especially Cracow and Kielce) are outliers, as they have generally high performing students. Nonetheless, results with the total sample of municipalities are also reported in the robustness section.

The Partitions borders under investigation were mostly set along rivers. Between the Prussian and Russian Empires it was drawn along the Drwęca and Prosna rivers (which are small waterways), whereas the half of the Austrian - Russian border was drawn along the Vistula river. Becker et al. (2014b) show that there are no significant differences between these regions in terms of geography and pre-Partitions historical characteristics. Similarly, Grosfeld and Zhuravskaya (2015), using the one-dimensional non-parametric regression discontinuity design, report only a small difference in altitude on the Austrian-Russian border. Yet, my own estimations in Table 4 Panel A show that there are also significant differences in temperature and precipitation on the Austrian-Russian border when the two-dimensional specification is used. The "jump" in altitude on the Russian side of the Austrian-Russian border is around 80 meters, precipitation is higher by 30mm and temperature drops by around 0.45 C°. The magnitudes are not large and they arise

[25] The borders of interest were established during the Congress of Vienna in 1815 after which they remained unchanged for almost 100 years.

[26] At the same time, almost the whole German population of these regions was expulsed to Germany.

most likely because of the riverbed of the Vistula. On the Prussian-Russian border the two-dimensional specification (Table 4 Panel C) also reports significant differences, but with smaller magnitudes.[27] Still, in order to control for the potential confounding effect, I control for the geographic and climate characteristics in the baseline regressions (the estimates are generally insensitive to their inclusion). Overall, I find it very unlikely that these natural differences could explain the educational differences between the borderlands or induce other dissimilarities in culture or institutions.

When the border exogeneity assumption hold, estimation of the discontinuous change in the outcome variable at the borders yields the causal effect of the Partitions of Poland. The channel of influence might be through social norms, migration or other process induced by the Partitions, for instance, urban settlement patterns.

3.4 Results

Figures 6 and 7 present relationship between the average student performance and distance to the Austrian-Russian and Prussian-Russian borders respectively. A drop in the 6th and 9th grade exam score can be seen clearly at the border between Austria and Russia (positive distance) indicating a strong and positive effect of the former Austrian Empire. Contrary to this, no visible effect can be seen on the Prussian-Russian border.

In Table 5 I report the coefficients and standard errors for the baseline model (with the quadratic polynomial and 50km bandwidth). Panels A and C show the two-dimensional specification. Columns (1) to (4) present regressions with the 6th grade low-stake exam score as a dependent variable, while columns (5) to (8) with the mathematics and science 9th grade high-stake exam score. The results for the rural sample are reported in columns (1), (2), (5), (6). Additionally, in columns (2), (4), (6), (8) I control for the set of geographic control variables.

Panel A presents the results for 301 rural municipalities located around the former Russian-Austrian border. Students living in the former Austrian partition outperform students from the former Russian side of the border on the 6th grade exam by on average 0.62 of standard deviation (σ)[28] and on the 9th grade exam by 0.42σ (columns (1) and (5)). The estimates drop to 0.54σ and 0.4σ respectively, once I add the set of geographic control variables ((2) and (6)). All the coefficients are strongly significant. The magnitudes and economic importance of the results are only slightly smaller than the Black vs. White achievement gap in the US in math for 8th graders is estimated to be around 0.88σ (Lee, Grigg and Dion, 2007). The smaller effects on the 9th grade high-stake exam is consistent with the favored social norm hypothesis, which predicts that the gap between regions with different social norms widen when there are no intrinsic incentives for obtaining a good score (see Section 4.1).

Similarly, Panel C depicts the same set of regressions for 206 municipalities from the former Russian-Prussian border. The coefficients are much smaller in absolute terms

[27] Consistently with Grosfeld and Zhuravskaya (2015), the one-dimensional specification in Table 4 Panel B and Panel D does not produce significant differences on neither of the borders.

[28] To obtain the effect of the Austrian or the Prussian Empires one has to simply change the sign of the coefficients reported in Table 5.

and are all insignificant. The estimates of the effect of the *Prussian* Empire, for the 6th grade exam (9th grade) are 0.03σ (0.07σ), and 0.06σ (0.13σ) when the geographic controls are included. Contrary to the Austrian-Russian border, these results show that students from the former Prussian zone do not perform better than those from the former Russian territories. In fact, the estimated absolute effects of the Austrian Empire on the Russian-Austrian border are significantly larger (at the 0.1% level) from the effects of the Prussian Empire on the Russian-Prussian border (the comparison is not reported).

This pattern is also visible in Figures 3 and 4. They show the modern map of Polish municipalities, along with the predicted values from the two-dimensional regression of the standardized exam scores (Z-scores) from 2011, specified as in columns (3) and (7) of Table 5. Notably, the level of the predicted value is clearly discontinuous at the Russian-Austrian border, but the same is not true for the Prussian-Russian border.

3.5 Robustness

The population size of municipalities might be endogenous with respect to the Partitions of Poland and limiting the sample only to the rural areas introduces a sample selection bias.[29] Therefore, as a first robustness check, I estimate Equation (1) on the total sample and include a categorical variable indicating the population size of a municipality. This is a less preferable sample, since large cities have generally better student performance and they might by chance significantly improve the average performance of the partitions. Nevertheless, as Table 5 columns (3), (4), (7) and (8) show, the results are practically insensitive to the inclusion of the urban areas.

A two dimensional polynomial is a natural way to model the relation between location and the outcome. However, Dell (2010) argues that the multidimensional regression discontinuity design might lead to an over-fit of a model at a discontinuity. On that account, I also run an one dimensional model, where $f(location_i)$ from Equation (1) is a polynomial in distance to either the Russian-Prussian or Russian-Austrian borders. I allow this polynomial to have different coefficients for the two sides of the borders.[30] I center the distance at the borders and define it such that on the Prussian or Austrian sides it is negative and on the Russian side positive. Panels B and D of Table 5 show the results. For the Austrian-Russian border, the magnitudes are smaller in absolute terms and in the case of 9th grade score they also lose significance. For the Prussian-Russian border, the magnitudes increase in absolute terms but they are still insignificant (with an exception of column (1)).

The baseline results might be sensitive to the specification choices. Table A4 reports estimates of the Partitions effects for different polynomials in latitude and longitude, along with different bandwidth choices. All regressions include the geographic controls.

[29] Suppose that the Austrian Empire positively affected the urban population growth. Limiting the sample only to municipalities smaller than 50 thousands people from both the former Austrian and Russian partitions, means that I compare "normal" municipalities from the Russian side with relatively disadvantaged ones from the Austrian side.

[30] This can be done by an inclusion of the interaction term between the partition dummy and the polynomial.

For the Austrian-Russian border, the results consistently show highly significant and positive effects of the Austrian empire on student performance. The effect varies from 0.53σ to 0.67σ in the case of the 6th grade exam and from 0.31σ to 0.48σ in the case of the 9th grade exam. Contrary to this pattern, the estimates of the Partitions effect on the Prussian-Russian border vary a lot across specifications. Importantly, the sign changes once the bandwidth is increased to 75km and 100km - which indicate that students living in the Russian zone are, in fact, perform better than those from the Prussian one. These contradict the findings from Table 5. Nevertheless, in most of the cases the coefficients are not significant.

The same set of specification choices is examined with the total sample (Table A5) and with the one-dimensional regression discontinuity design for the rural sample (Table A6). Both tables consistently show highly significant (except cubic and quartile specifications in Table A6) and positive effects of the Austrian Empire. Results for the Prussian Empire are similar as previously.

In order to see whether the results are driven by a particular year, I estimate the baseline Equation (1) by OLS for each year separately and the rural sample. Figure 10 depicts the estimated Partitions effects for each year-sample, along with the 95% confidence intervals. Similarly as in the pooled sample, the effect of the Austrian Empire is consistently positive and significant in the case of the 6th grade low stake exam and shifted toward zero in the case of the 9th grade high stake exam. By contrast, the effect of the Prussian Empire is null across years and types of the exam.

Next, I check whether the results are sensitive to an inclusion of a set of time-variant modern controls. The variables reflect general educational differences (local government expenditures per capita on education, kindergarten and secondary school attendance) and the socio-economic conditions (local government total expenditures per capita, unemployment ratio, population level, population density and migration balance). Table A1 provides the exact definition of the variables. Importantly, note that these covariates are endogenous, that is, they are also affected by the Partitions of Poland (see Section 4.3 and Table 6). As such, the augmented regression "switches off" some potential channels of influence of the Partitions of Poland on outcomes and leads to "bad control bias" (Angrist and Pischke, 2008).[31] Indeed, as reported in Table A7, the effect of the Austrian partition drops slightly, but remains significant and economically important. Depending on polynomial and bandwidth, the estimated coefficients range from 0.36σ to 0.45σ in the case of the 6th grade exam and from 0.2σ to 0.32σ in the case of the 9th grade. On the other hand, similarly as previously, the estimates of the effect of the Prussian partition are either not significant or have an opposite sign.

The current border between the vovoidships (NUTS2 administration level) overlaps almost completely with the former Russian-Austrian border. If vovoidships influence the quality of education, their effect could be mistakenly confounded with the effect of the Austrian or Russian partition. There are two arguments against this possibility. Firstly, the Polish education system is considered very decentralized (Herbst, Herczynski and Levitas, 2009). A local municipality's government manages the school network of almost

[31] However, the direction of the bias is not clear.

all public elementary and lower secondary schools, at the same time the role of the central government is limited to financing education and enacting general resolutions. The vovoidship administration is thus practically irrelevant for the educational governance. Consistently, Table 6 shows that there is no significant effect of the Austrian Empire in municipality's educational spending per capita. Secondly, since the former Russian-Prussian border does not overlap with the administrative borders, I can include the vovoidship dummies in Equation (1). It turns out that none of these dummies are significant, which suggests that the vovoidship administration is not relevant for the performance of students. This observation is also consistent with Herbst (2004).

Finally, I run a set of placebo experiments in which step-by-step I artificially move the Austrian-Russian border by 5km to the North or to the South (at most around 100 km). I run the baseline specification, with the artificial borders, and define the "Russian" dummy as an area North from the artificial borders. Figure 11 Panel B presents Z-tests[32] of the placebo Partitions effects, for each artificial border. Notably, only the actual border (at point 0) is an outlier. Analogously, I move step-by-step the Prussian-Russian border by 5km to the West or to the East (at most around 100km) and define the "Russian" dummy as an area East from the artificial borders. Figure 11 Panel B shows Z-tests. This time the actual border is not different from the other artificial borders.

Taken together, these results show that the former Austrian Empire has a positive effect on the exam scores once compared with the Russian Empire. The effect is stable across specification, highly significant and large. Conversely, the effect of the former Prussian Empire is usually insignificant, low and changes sign across specifications.

4 Channels of Persistence

In this section I use modern data to investigate the channels of persistence. There are many possible explanations of the observed pattern in exam score. I highlight the social norms channel, which other studies suggest to be the most important.[33] In particular, the existing literature on the long-lasting effects of the Central and Eastern European Empires underlines the importance of inter-generational transmission of norms and values. Grosfeld and Zhuravskaya (2015) provide evidence that the Partitions of Poland has exerted a long lasting effect on religiosity and belief in democratic values through the inter-generational within-family transmission of social norms. More broadly, Becker et al. (2014b) and Karaja (2013) show that the government polices introduced by the Habsburg and Ottiman Empires still affects trust toward local state and acceptance of corruption. My results show that people living in the former Austrian Empire have higher social norm with respect to the *local educational system*, even though they do not

[32] Z-test is defined as a ratio of an estimated coefficient and a corresponding robust standard error. This is an asymptotic analogue to the classic T-test. Z-test has an asymptotic Normal distribution.

[33] The studies not focused on the Central and Eastern Europe are for example Putnam et al. (1994); Akerlof and Kranton (2010); Cassar et al. (2013); Sakalli (2014); Feir (2015).

necessary have higher general social norm toward intelligence and higher (non-local) education. At the same time, I also discuss the alternative channels, in particular, current and historical skill-biased migration.

4.1 Social Norms toward Local Education System

I provide three pieces of empirical evidence to argue that the Partitions of Poland have created different social norms toward local educational system. Firstly, I develop a simple, *social norm*-based model of student schooling efforts, which is consistent with the observable pattern that the effect of the Austrian Empire is larger in the case of the low stake exam then in the case of the high stake one. Secondly, I use the data on proxies for social values to directly compare people's attitudes toward education across the former Partition borders. Finally, I show that the Austrian partition has a positive and large effect on kindergarten attendance, which cannot be explained by the historical supply of kindergartens.

4.1.1 The Low Stake vs. The High Stake Exams

Suppose that the test score T_{ig} of student i from grade g is a function of a student's effort e_{ig} (which summarizes a student's input into education), of other grade-invariant characteristics X_i and of an idiosyncratic shock ϵ_{ig}.

$$T_{ig} = \alpha + \beta e_{ig} + \gamma X_i + \epsilon_{ig}$$

In order to model the level of schooling effort I follow Akerlof and Kranton (2002). The authors propose a formulation of a student utility function, which combines the standard motivation of an individual (such as the direct costs and benefits of education) with the social norm based motivation, penalizing the individual for not copying with the existing social norm:

$$U(e) = p(we - \frac{1}{2}e^2) - (1-p)\frac{1}{2}(e-E)^2 \tag{2}$$

where e is the amount of schooling effort, w is wage rate per unit of effort, parameter p is weight given to pecuniary benefits and costs of effort, and E is the social reference point (social norm) with respect to the level of effort. In this formulation, the optimal choice of student's level of effort depends on general economic forces and social expectations.

Next, consider a minor modification of the Akerlof and Kranton (2002) student utility function, as defined in Equation (2). Suppose there are two regions R : former Austria and Russia, and that they differ with respect to the social norm toward schooling effort $E(R)$. There is also a *common* social norm toward the future earnings A, which can be arbitrarily large.

In the maximization problem for the 9th grade high-stake exam score, a student chooses a level of effort, which maximizes the following utility function:

$$U(e) = p(we_{i9} - \frac{1}{2}e_{i9}^2) - (1-p)(\frac{1}{2}(e_{i9} - E(R))^2 + \frac{1}{2}(we_{i9} - A)^2)$$

The optimal level of schooling effort is given by:

$$e_{i9}^* = \frac{pw + (1-p)(E(R) + wA)}{1 + (1-p)w^2}$$

Assuming that $E(Austria) > E(Russia)$, the average difference in a student's level of effort for the former Austria and Russia is:

$$GAP_9 = e_{AUS,9}^* - e_{RUS,9}^* = \frac{(1-p)(E(AUS) - E(RUS))}{1 + (1-p)w^2} > 0$$

In the case of the 6th grade low-stake exam score, the maximization problem is simpler. This exam score does not matter for the future educational career, therefore it will not have an impact on the future wages. The utility and the first order condition are thus:

$$U(e) = -p\frac{1}{2}e_{i6}^2 - (1-p)\frac{1}{2}(e_{i6} - E(R))^2$$

$$e_{i6}^* = (1-p)E(R)$$

The gap between the regions in the level of effort is then:

$$GAP_6 = e_{AUS,6}^* - e_{RUS,6}^* = (1-p)(E(AUS) - E(RUS)) > 0$$

Since $1 + (1-p)w^2 > 1$, it follows that $GAP_{6th} > GAP_{9th}$, as long as $p < 1$. Under the assumption that the exogenous students' characteristics are similar around the border,[34] we might conclude that:

$$\bar{T}_{Austria,6} - \bar{T}_{Russia,6} > \bar{T}_{Austria,9} - \bar{T}_{Russia,9} > 0$$

Which is consistent with the empirical results presented in Section 3, namely that the effect of the Austrian Empire is larger in the case of the Low Stake exam, than in the case of the High Stake one (Tables 5 - A7 and Figure 10). In the case of the Prussian-Russian border, there is no difference in social norms (i.e., $E(Prussia) = E(Russia)$), so there is no gap in the performance of students.

4.1.2 Proxies for Social Norms

I use survey data on attitudes toward education to see whether current social norms systematically differ across the historical borders of the Partitions of Poland. The primary source of the data is the two waves (2011 and 2013) of the Social Diagnosis survey, which include over 45,000 individuals from almost all counties in Poland. The data does not permit to directly measure social norms with respect to local schools. Nevertheless,

[34] Which is a stronger assumption than in the case of the regression discontinuity design.

the questionnaire asks, whether education is important for a good life, whether a respondents are satisfied with received education and their desired level of education for their children. The variables are described in Table A3. A second source, with smaller sample size, is the Life in Transition Survey (LiTS, organized by the European Bank of Reconstruction and Development), which includes around 7000 individuals from 350 primary sampling units (PSU) from Poland and asks questions about first or second priority of education in governmental spending, private expenditures on education and opinion about the role of intelligence and skills in life success. A third source is the Educational Value Added survey (EVA, conducted by the Educational Research Institute in Warsaw), which includes approximately 10,000 parents and asks about the role of family tradition in school selection. For the detailed descriptions of variables from LiTS and EVA see Table A2.

The Social Diagnosis survey reached respondents from almost all counties in Poland, therefore I can use a geographical regression discontinuity design, where location of a respondent i is determined by location of her county c. The estimated equation is:

$$y_{ic} = \alpha + f(location_c) + \beta D_c + \gamma G_{ic} + \delta X_c + \kappa_w + \epsilon_{ic} \tag{3}$$

where notation is similar as in Equation (1). Additionally, I control for a set of the county-level socio-economic characteristics X_c, which are described in Table A1. The observations come from the two waves, but these are for different respondents, therefore I pool the sample and include survey wave fixed effects κ_w. Since location and assignment to a partition vary by counties (which are higher administration unit than municipalities), it is important to assure that there is enough variation in location and that the model is not over-fitted. For these reasons, I use the total sample, along with 100km bandwidth and a one dimensional polynomial in distance. Depending on the outcome variable, I use either Probit or Ordered Logit estimators.

The empirical strategy for the LiTS and EVA surveys is similar, except that I need to account for the fact that the sample is based on only several PSUs that are located near the former borders of the Partitions. I use Equation (3), with the whole sample and 100km bandwidth, but in this case I do not include a polynomial in location or distance.[35] Beyond this limitation, one also has to keep in mind that the LiTS and EVA samples are not representative at the regional level. Depending on the outcome variable, I use OLS, Probit or Ordered Logit estimators.

Table 7 (8) presents estimates of the effect of the Partitions of Poland using the sample from Social Diagnosis (LiTS and EVA) and the reported numbers are the marginal effects at the border (the average marginal effects). The results show that people living in the former Austrian partition are around 5% more likely to say that education is important for a good life (Table 7, columns (1) and (2)) and around 19% more likely to say that public education should be given first or second priority in governmental spending (Table 8, column (2), but (1) is insignificant). They are also 6% more likely to claim that family tradition is important in their local school choice ((Table 8, column

[35]The reason for not using a regression discontinuity is that there are just few locations around the borders and one might over-fit the model.

(7), but (8) is insignificant). However, at the same time, they are around 20% *less* likely to desire higher education for their children (Table 7, columns (5) and (6)) and 20% *less* likely to agree that intelligence and skills are important in life success (Table 8, columns (3) and (4)). Finally, the sign of insignificant estimates suggests that people from the Austrian partition are more satisfied from received education ((Table 7, columns (3) and (4)), but spend less on education of their children ((Table 8, columns (5) and (6)).

These results can be interpreted as an evidence for a positive social norm toward the *local* educational institutions. People from the former Austrian Empire perceive education as important in their life, want more funds directed to public schools[36] and underscore the long run relationships of their families with local schools.[37] On the other hand, their attachment to the local schools might lead to a lesser trust toward the non-local ones. This would explain why people are less desirable to send their kids to an institution of higher education, which is usually outside the local environment. Nevertheless, they also perceive intelligence and skills as less important and, even though this question is outside the educational context, this is a clearly puzzling result. Overall, in the light of the model outlined in the previous section, one could argue that people living in the former Austrian empire have higher social norms toward schooling effort in local school, that is $E(Austria) > E(Russia)$, but lower toward the expected future career and earnings $A(Austria) < A(Russia)$.

In the case of the Prussian-Russian border, the only significant estimates show that people from the Prussian partition are less likely to agree that education is important for a good life, but more likely to agree that intelligence/skills are important in life success. These would suggest that there is also a small difference in social norms, that is, $E(Prussia) < E(Russia)$ and $A(Prussia) > A(Russia)$, however, these seem to be too small to produce strong and systematic differences in the performance of students.

4.1.3 Kindergarten Attendance

The third piece of evidence shows that the Austrian partition has a positive and significant effect on kindergarten attendance. Table 9 presents the estimates of the border discontinuities in kindergarten attendance, defined as a share of children aged 3-5 who attend an institution of pre-education. I use the regressions specified as in Equation (1).[38] The results show that the kindergarten attendance ratio is higher on the Aus-

[36] The stronger belief about first or second priority of education in a governmental spending in the Austrian partition can be alternatively explained by a poor quality of a local public education. However, if this would be true we would rather observe a negative impact of the Austrian Empire on the exam scores. As the previous section shows, this is not the case. Moreover, there is no systematic difference between the Austrian and Russian Partition in terms of a school infrastructure, as reported in Herczynski and Sobotka (2013), and there is no visible effect of the Partitions of Poland on the estimates of school value added (see Table 6).

[37] Interestingly, parents from the Austrian Partition are also 5% more likely to *agree* that corporal punishment is important for a child development (Table 7, column (7)). This suggests that they are also more conservative than parents from the Russian partition.

[38] Variables are at the municipality level. In all the regressions I use 50km bandwidth and quadratic polynomial in longitude and latitude. I pool years and use the Random Effect estimator.

trian side by 3-7 percentage points compare to Russia. On the other hand, there is no difference on the Prussian-Russian border.

The effect could be a result of the historically determined demand for pre-education or the historical supply of institutions of pre-education. This later explanation implies that the higher kindergarten attendance in the former Austrian partition is due to the inherited buildings/institutions from the 19th century. Unfortunately, no data exist to test this hypothesis. Nevertheless, if this would be true we would also observe a positive impact of the Prussian Empire as the historical school network was denser in the Prussian partition (see Table 2). As this section shows, this is not the case.

On the other hand, historically determined demand for pre-education is consistent with the social-norm channel. Pre-education is not obligatory in Poland and it is partially determined by parent's willingness to send their children to kindergartens. The higher social norm toward local educational system, which also applies to local kindergartens, can thus explain the positive effect of the Austrian Empire on kindergarten attendance.

4.2 Migration

Modern skill-biased migration might be an alternative explanation of the observed effect of the Partitions of Poland. If, for instance, only high achievers migrate to the former Austrian partition and only low-achievers to the former Russian partition, one should expect to find a significant gap in the average student performance.[39] In order to evaluate this possibility, I follow Dell (2010) and exploit the student-level data on the exam scores (from the Central Examination Board) and adjust it for the municipality-level data on the share of in-migrants (from the Central Statistical Office of Poland).[40] Specifically, I assume the "worst" migration scenario outlined above, and on the former Austrian lands I trim the *top* of the distribution of the student exam scores according to the share of in-comers at the municipality level. Analogously, on the former Russian lands I trim the *bottom* of the distribution. Next, I aggregate the trimmed data to the municipality level and repeat the estimations from Section 3.4 (the baseline specification of the Equation (1)). Consistently with expectations, the effect of the Austrian partition for the 6th (9th) grade score drops from 0.62σ to 0.47σ (0.44σ to 0.28σ), but it still remains highly significant and economically relevant.[41] These suggest that the modern migration itself is unlikely to explain the observed effect of the Partitions.

Nevertheless, the effect might be affected by historical migration, if the selection of migrants is not orthogonal to the current performance of students. My main results could be, for instance, explained by migration of high skilled people from the Russian to the Austrian part (or low skilled vice versa) and/or migration of low skilled people from the Austrian part to third countries. Existing qualitative evidence suggests that

[39] Please note that the potential reasons for the migration might be endogenous with respect to the Partitions of Poland.

[40] For each year I adjust the student level data using information on share of in-migrants from that year.

[41] In a specification with the geographic control variables, the effect drops from 0.55σ to 0.41σ (0.42σ to 0.26σ). Full results available upon request.

the first possibility is unlikely. Labuda (1971) argues that majority of migrants between the partitions were seasonal workers and they did not settle permanently.[42] Moreover, the economic situation in Galicia was the hardest and the level of industrialization the lowest, so there were no strong incentives for skilled workers to migrate there. On the other hand, migration to third-countries is harder to assess, as there are no clear insights in the literature about the 19th century emigrants' skills from the partitioned Poland. Zubrzycki (1953) hypothesizes that migrants from the Grand Duchy of Poznań were more likely to be better educated than migrants from Galicia and the Congress Kingdom.[43] However, as reported in Abramitzky, Boustan and Eriksson (2012), the late 19th century U.S. immigrants from Austria and Prussia had wages similar to the US population, even though there was a substantial variation in wages across immigrants coming from other European countries.

GUS (2003) and Zubrzycki (1953) offer limited aggregated data about migration and population characteristics from the 19th and early 20th century. Table 10 Panel A presents migration balance and Panel B presents migration as the share of 1910 population for each partition. The numbers show a large outflow of population from the Austrian and Prussian part. Importantly, however, the Austrian partition does not seem to be unique and, in fact, migration was larger in the Prussian partition. Similar picture emerges from the data on general population characteristics (Panel C and D). The population trends and age structures are similar across the partitions and the Austrian part does not show any anomalies, that could result from some unique migration pattern, for instance, over or under representation of the middle-age population groups.

4.3 Other Channels

Alternative channels of persistence of the Partitions of Poland could be based on other differences between the Empires, such as urbanization and economic policies or the expansion of universities. In order to shed light on these, I estimate the border discontinuities of various socio-economic characteristics described in Table A1. I use the regressions specified as in Equation (1),[44] but with the dependent variables transformed to the natural logarithms. Note that, similarly as the exam scores, these variables are endogenous, and might reflect the effect of the historical education systems or some other channels.

Table 6 column (1) reports estimates of the effect (semi-elasticity) of the Russian Empire on the Prussian-Russian border, while column (2) on the Austrian-Russian border - therefore changing the sign yields either the effect of the Prussian or Austrian Empire.

[42] However, because of the universities, there was a small migration of students to the Austrian part but it was limited to major cities i.e., Cracow and Lviv (Cohen, 1996).

[43] Nevertheless, given the size of migration (see Table 10), the majority of migrants had to be relatively uneducated.

[44] Variables are at the municipality level. In all the regressions I use 50km bandwidth and quadratic polynomial in longitude and latitude, the sample is limited to the rural areas, and only the partition dummy and geographic controls are included as independent variables. For the time-varying variables I use the Random Effect estimator, for the time-invariant variables I use OLS.

The effect of the Partitions on the Prussian-Russian border is insignificant, except for the share of people in agriculture and class size.

The picture looks different in the case of the Austrian-Russian border. Firstly, the Austrian Empire positively affects the demographic characteristics, namely the level of population, density, and migration balance. This could be an alternative channel if different urbanization patterns emerged during the Partitions period and urbanization influenced the quality of education. For instance, population density could affect class size, which in turn influences student performance (Angrist and Lavy, 1999).[45] However, there is no difference in class size between the former Austrian and Russian lands. Moreover, the reported estimations in Section 3.5 include the demographic characteristics as control variables. If urbanization patterns are the main channel of persistence, I should find an insignificant and small effect of the Austrian partition. Contrary to this, as reported in Table A7, the positive and strong effect of the Austrian partition is insensitive to the inclusion of the demographic characteristics.

Secondly, the economic situation on both parts of the former border is similar, except for a higher unemployment rate on the Austrian side. That being the case, it is unlikely that general economic forces, such as returns to education, could be a driving mechanism. Finally, there is a significant positive effect on the share of people with higher education. This is consistent with the highlighted hypothesis, as one would expect that the social norms toward education are influencing not only performance of students on the exams scores, but also other educational outcomes.

5 Identity as a Determinant of Persistence

This section argues that the Austrian Empire succeeded in creating a positive social norm toward education, because of a positive interaction between institutional quality and identity. Consistent with this hypothesis, Steele and Aronson (1995) and Akerlof and Kranton (2002) provide theoretical and empirical evidence that identity is associated with social norms affecting an individual's schooling choices, school-student relationships and student achievements. A similar hypothesis is also explored by Sakalli (2014), who documents that the Muslim identity of the Turks has been reinforced by the past coexistence with the Armenians, which in turn, has changed the long-run social/cultural norm toward the secular education.

In the first part of this section I conceptualize the hypothesis, in the second I provide a suggestive evidence for it. In particular, I use historical data to measure the within-partition variation in the 19th century educational attainment and link it with the current-day performance of students.

[45] But this scenario would rather imply that classes are larger on the former Austrian lands, and so the performance of students *lower*.

5.1 Conceptualization

Suppose there are two time periods (i.e., the 19th century and the modern time). In the first period, an educational system is introduced and individuals decide how much schooling effort e_1 they should exert.

$$U(e_1) = w_1 e_1 - \frac{1}{2} a_1 e_1^2 \tag{4}$$

Utility comes from a difference between the benefits and costs of education. These are determined by schooling effort, wage premium w_1 and the cost parameter a_1. The optimal level of schooling efforts in the first period is a fraction of the wage premium over the cost parameter.

$$e_1^* = \frac{w_1}{a_1}$$

It can be argued that formal institutions, that are inconsistent with ethnic identity, will impose higher costs for an individual. In the case of the Partitions of Poland, the Prussian education system required Polish students to learn in German and to study anti-Polish material. The cost of education included then an additional cost of learning a foreign language and an intrinsic discomfort coming from the ethnic intolerance. Conversely, the Austrian system offered similar institutions, but with the Polish language of instruction and without the anti-Polish curriculum. Since the returns to education were relatively modest at that time (Cvrcek and Zajicek, 2013) the model would imply that the relatively lower costs of education in the Austrian Empire contribute to the higher schooling effort.

Next, suppose that a social norm about some behavior emerges within a society when all individuals are consistently finding this behavior as optimal. The social norm can be then transmitted through generations and still affect a student's optimal choice of schooling effort (Bisin and Verdier, 2001; Patacchini and Zenou, 2011; Spolaore and Wacziarg, 2014). The difference for the future generations is that the past institutions affect individual's optimal behavior not through the standard part of the utility function, but through the social norm part. Therefore, in the second period, individuals are also facing a social norm with respect to schooling effort. Using the utility function defined in Equation 2 and assuming for simplicity that the wage premium and the cost parameters in the second period are equal to unity ($w_2 = a_2 = 1$), we have:

$$U(e_2) = p(e_2 - \frac{1}{2} e_2^2) - (1-p)\frac{1}{2}(e_2 - \frac{w_1}{a_1})^2 \tag{5}$$

Note that the social norm with respect to the schooling effort is the optimal level of schooling effort from the previous period. An individual's choice in the second period depends then positively on the past wage premium and negatively on the past cost parameter

$$e_2^* = \frac{p + (1-p)(\frac{w_1}{a_1})}{1 + (1-p)}$$

$$\frac{\partial e_2^*}{\partial a_1} = -\frac{(1-p)(\frac{w_1}{a_1^2})}{1+(1-p)} < 0$$

In other words, this simple model implies that the relatively lower costs of education in the Austrian Empire became a crucial factor for the formation of the social norm and thus for the future schooling effort.

5.2 Evidence

The analysis so far assumed that the effect of the Austrian and the Prussian Empires is the same across municipalities from the same partition. In this subsection I relax this approach and exploit the county (deanery) - level historical data on educational attainment, to measure the *within* partition variation in the exposure to the 19th century institutions, and link it with the current-day performance of students. If the hypothesis is true, one should observe a positive causal effect of the past educational attainment, measured by the elementary school enrollment ratio, on the current-day quality of education in the former Austrian Empire, but a null or negative effect in the former Prussian Empire.

The historical data on educational characteristics come from the 19th century censuses. In the case of the Prussian Empire, the source is the Ifo Prussian Economic History Database (Becker, Cinnirella, Hornung and Woessmann, 2014a). For the Austrian Empire, I use the data collected by Cvrcek and Zajicek (2013). Unfortunately, no such data is available for the Russian Empire. The data for the Prussian partition is based on the 19th century Prussian counties, which I assigned to modern municipalities using GIS methods and maps provided by Kashin and Ziblatt (2012). The data for the Austrian part is based on the 19th century deaneries, which is an administration unit of the Catholic Church. As there is no GIS map of deaneries from the Austrian Empire, I manually matched modern municipalities with their historical deaneries using information from Dobrowolski (1886) and the geographic dictionary by Sulimierski, Chlebowski and Walewski (1895). Unfortunately, for the Austrian part, only the census of 1865 offers data disaggregated to the deanery level. In order to keep a comparable time frame, I therefore use only data from the Prussian census of 1864. The variable of interest is the total educational attainment at the obligatory, elementary education level, which is defined as percent of children enrolled in elementary school (both public and private). In order to ensure comparability across the censuses, I standardized the measures, so that for each partition they have mean of zero and standard deviation of one.

I first document a simple correlation of the standardized 19th century educational attainment and the average of the modern exam scores (2005-2011). Figure 12 presents correlation for the 6th grade low stake exam score and Figure 13 for the the 9th grade high stake exam, in breakdown by the rural and total samples. Consistently with the hypothesis, we can observe positive correlations in the case of the Austria partition (which is significant for the 9th grade exam), but negative in the case of the Prussian

partition (significant for the 6th grade exam and the rural sample).

The reported correlations might possibly reflect the omitted variable bias. For instance, a favorable location of a municipality might affect its long run prosperity and influence historical educational attainment along with the current performance of students. In order to limit the bias, I run regressions of the standardized 6th or 9th grade exam scores on the historical educational attainment, and control for a quadratic polynomial of longitude and latitude, and geographic characteristics. I pool the data from the Austrian and Prussian partitions, include dummy for Austria and interact it with the historical measure:

$$y_{mcpt} = \alpha + \beta_1 A_{cp} + \beta_2 Aus_p + \beta_3 Aus_p A_{cp} + \gamma G_{mcp} + \delta X_{mcpt} + \epsilon_{mcpt} \tag{6}$$

where y_{mcwt} is the outcome variable for municipality m from county (deanery) c and partition p at time t (available for 2005-2011). A_{cp} is the educational attainment from the 1860's, which is available at the county (deanery) level, Aus_p is a dummy for the former Austrian partition, G_{mcp} is a set of exogenous geographic controls. In addition to this, in some specifications I include a set of time variant municipality socio-economic characteristics X_{mcpt}, defined in Table A1. I pool all the years and use the Random Effect estimator. The standard errors are clustered at the county (deanery) level.

Table 10 presents the estimates of Equation (6). Columns (1) and (2) show that the correlations between the educational attainment and the 6th grade exam scores are significant and negative for the former Prussian partitions. One standard deviation increase in the attainment is connected with 0.14 - 0.26σ decrease in the student performance. Importantly, this correlations become close to zero or positive for the former Austrian partition. The correlations in the case of the 9th grade exam (columns (3) and (4)) have the same sign, but the coefficients are smaller in absolute terms and insignificant.

The simple control on observable approach is unlikely to solve the endogeneity problem. However, assuming that the remaining bias is the same in both regions, the historical expansion of the education system has more positive effect on the current student performance in the former Austria than in the former Prussia. This is in line with the proposed hypothesis, as in the Austrian Empire there was a *positive* interaction between identity and institutions. Hence, the social norms affecting student performance may have been more likely to emerge in municipalities with a larger attachment to the historical Austrian educational system. Alternatively, because of the *negative* interaction between the institutional quality and identity, the more intensive historical exposure to the Prussian education might lead to a stronger opposing social norm toward the educational system. This norm leads to a lower schooling effort and thus decreases the performance of students. Furthermore, using arguments from Section 4.1.1, the weaker relationship in the case of the 9th grade high stake exam points to the importance of the social norms as a channel of persistence.

Nevertheless, distinctive characteristics of the Austrian education system, other than the positive interaction with identity, might be another source of persistence. Especially important seem to be inclusiveness of the Austrian *secondary* education and the existence

of two universities and one technical college in Galicia.[46]

6 Conclusions

This paper argues that the Partitions of Poland provide a unique natural experiment for studying the determinants of institutional persistence. First, I exploit this setting to investigate the long lasting effect of the 19th century educational systems, which were imposed by Austria, Russia, and Prussia, on modern educational outcomes. Despite the modern similarities of the former borderlands of the Empires, I estimate a positive and large effect of the former Austrian Empire compared to the former Russian Empire, but no effect of the Prussian Empire compared to the Russian. The magnitude of the effect of the former Austrian Empire is similar to the Black vs. White achievement gap in the US.

How can we explain these results, and what can we learn about persistence of institutions? The main hypothesis argues that an interaction between institutional quality and identity might be crucial for the creation of a positive social norm toward institutions, and thus for a long lasting persistence. In particular, because the Austrian education system was actively supporting Polish identity, positive norms toward education system were more likely to emerge in the Austrian partition and these could be transmitted through generations and still affect student and parental effort. The Prussian partition serves as the counter factual situation, where almost identical educational system is used as a tool of Germanisation, and no positive social norm affects modern performance of students.

This result might be of a crucial important for policymakers who wish to improve the situation of permanently underdeveloped regions. For instance, the implication of my hypothesis is that good educational institutions are more likely to affect long run development if they are not in opposition to the social identity of a population of interest. One might consider provision of schooling in a minority's language (i.e. for Hungarians in Romania, Poles in Lithuania or Russians in Ukraine) as an example of such policy. On the other hand, this paper suggests that large interventions aimed at equalizing educational differences, as carried out by the post-WWII communist government in Poland, might have a large effect on quantity of education, but limited on quality of education.

The proposed sources of persistence can also shed light on the existing, and often puzzling, findings in the literature on institutional persistence. For instance, a study of Africa, by Michalopoulos and Papaioannou (2013), reports that the pre-colonial ethnic institutions matter for the long run development to a much larger extent than the national ones created by the colonial powers. Similarly, in the Indian context, Iyer (2010) reports a negative effect of the British colonial rules and positive of the native states. These are consistent with my hypothesis that institutional persistence is determined by the interaction between institutions and identity. The native institutions were to a larger extent compatible with the existing social identities. In contrast, the external

[46]There were no institutions of higher education in the Prussian part of Poland.

powers imposed the national institutions based on the borders, which broadly ignored the ethnic division.

The interaction between institutions and ethnic identity can be perceived as an ingredient of the interethnic inequalities and modes of cooperation. As suggested by Jha (2013), a limited ethnic assimilation might lead to "the presence of nonreplicable and nonexpropriable source of interethnic complementarity", which fosters ethnic tolerance by increasing the long run cost of potential ethnic conflict. In addition, Alesina and La Ferrara (2005) claim that the outcome of interethnic complementarity is higher specialization in an economy, which, in turn, increases productivity.[47] In the context of Native American tribes, Dippel (2014) presents evidence that forced integration of linguistically homogeneous sub-tribal groups has a negative effect on the long run economic development through the quality of local governance. I add another channel to this debate by pointing out that a lack of (forced) assimilation could ensure that all ethnic groups share the long run benefits from institutional change.[48]

Finally, this paper contributes to the literature on the formation of human capital, by pointing out the importance of social norms for the quality of education. This is of increasing importance for the developed countries, where the existing compulsory schooling law ensures that all children have right to a free and public education, and the quality of education became a major determinant of a country's educational success (as measured for example by the PISA score ranking). Moreover, many authors have pointed out that proper institutional design is crucial for the formation of human capital (Galor and Moav, 2006; Becker and Woessmann, 2009; Goldin and Katz, 2009; Cantoni and Yuchtman, 2013). However, this paper suggests that analyzing institutions without a social context might be insufficient.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson, "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," Technical Report, National Bureau of Economic Research 2012.

Acemoglu, Daron and James A Robinson, "Persistence of Power, Elites and Institutions," *The American Economic Review*, 2008, *98* (1), 267–293.

_ , Simon Johnson, and James A Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *The American Economic Review*, 2001, *91* (5), 1369–1401.

[47] On the other hand, because of heterogeneity of preferences, fragmented societies are more likely to have inefficient and poorly managed public goods (La Porta, Lopez-de Silanes, Shleifer and Vishny, 1999). Also, the existence of heterogeneous ethnic groups, which are clearly distinguishable, might motivate the ruling party to use ethnic conflict as a tool of expropriation (Caselli and Coleman, 2013) or prevent voters from replacing an inefficient politician (i Miquel, 2007).

[48] In the case of education, one might further argue that the institutional change leads to more educated society, which likely increases ethnic tolerance.

Akerlof, George and Rachel Kranton, "Identity and schooling: Some lessons for the economics of education," *Journal of Economic Literature*, 2002, pp. 1167–1201.

_ and _ , *Identity economics: How our identities shape our work, wages, and well-being*, Princeton University Press, 2010.

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance.," *Journal of Economic Literature*, 2005, *43* (3), 762–800.

Angrist, Joshua D and Jorn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press, 2008.

Angrist, Joshua D. and Victor Lavy, "Using Maimonides' Rule To Estimate The Effect Of Class Size On Scholastic Achievement," *The Quarterly Journal of Economics*, May 1999, *114* (2), 533–575.

Banerjee, Abhijit and Lakshmi Iyer, "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *The American Economic Review*, 2005, *95* (4), pp. 1190–1213.

Basten, Christoph and Frank Betz, "Beyond Work Ethic: Religion, Individual, and Political Preferences.," *American Economic Journal: Economic Policy*, 2013, *5* (3), 67–91.

Becker, Sascha O and Ludger Woessmann, "Was Weber wrong? A human capital theory of Protestant economic history," *The Quarterly Journal of Economics*, 2009, *124* (2), 531–596.

_ , Francesco Cinnirella, Erik Hornung, and Ludger Woessmann, "iPEHD—The ifo Prussian Economic History Database," *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 2014, *47* (2), 57–66.

_ , Katrin Boeckh, Christa Hainz, and Ludger Woessmann, "The Empire Is Dead, Long Live the Empire! Long-Run Persistence of Trust and Corruption in the Bureaucracy," *The Economic Journal*, 2014.

Bisin, Alberto and Thierry Verdier, "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic Theory*, 2001, *97* (2), 298–319.

Bowen, James, *Soviet education: Anton Makarenko and the years of experiment*, University of Wisconsin Press, 1962.

Burke, Peter, *The New Cambridge Modern History: Volume 13, Companion Volume*, Vol. 13, Cambridge University Press, 1979.

Cantoni, Davide and Noam Yuchtman, "The political economy of educational content and development: Lessons from history," *Journal of Development Economics*, 2013, *104*, 233–244.

Caselli, Francesco and Wilbur John Coleman, "On the theory of ethnic conflict," *Journal of the European Economic Association*, 2013, *11* (s1), 161–192.

Cassar, Alessandra, Giovanna d'Adda, and Pauline Grosjean, "Institutional Quality, Culture, and Norms of Cooperation: Evidence from a Behavioral Field Experiment," *UNSW Australian School of Business Research Paper*, 2013, (2013-10).

Chubarov, Alexander, *The fragile empire: A history of imperial Russia*, Blackwell Publisher, 2000.

Chuminski, Jedrzej, "Gospodarcze i spoleczne skutki zaborow Polski," 2008.

Cinnirella, Francesco and Ruth M. Schueler, "The Price of Decentralization: Linguistic Polarization and the Provision of Schooling," *mimeo*, 2015.

Cohen, Gary B, *Education and middle-class society in imperial Austria, 1848-1918*, Purdue University Press West Lafayette, IN, 1996.

Cvrcek, Tomas and Miroslav Zajicek, "School, what is it good for? Useful Human Capital and the History of Public Education in Central Europe," Technical Report, National Bureau of Economic Research 2013.

Davies, Norman, *God's Playground A History of Poland: Volume I: The Origins to 1795*, Vol. 1, Oxford University Press, 2005.

_ , *God's Playground A History of Poland: Volume II: 1795 to the Present*, Vol. 2, Oxford University Press, 2005.

Dell, Melissa, "The persistent effects of Peru's mining mita," *Econometrica*, 2010, *78* (6), 1863–1903.

_ **, Nathan Lane, and Pablo Querubin**, "State Capacity, Local Governance, and Economic Development in Vietnam," *Technical Report*, 2015.

Dippel, Christian, "Forced coexistence and economic development: evidence from Native American Reservations," *Econometrica*, 2014, *82* (6), 2131–2165.

Dobrowolski, Tadeusz Nieczuja, "Biskupstwo Tarnowskie. Rys Historyczny," *Biblioteka Warszawska*, 1886, *3*, 365–373.

Feir, Donna, "The Intergenerational Effect of Forcible Assimilation Policy on Education," *Technical Report*, 2015.

Galor, Oded and Omer Moav, "Das human-kapital: A theory of the demise of the class structure," *The Review of Economic Studies*, 2006, *73* (1), 85–117.

Gillis, John R, *The Prussian Bureaucracy in Crisis, 1840-1860: Origins of an Administrative Ethos*, Stanford University Press, 1971.

Goldin, Claudia Dale and Lawrence F Katz, *The race between education and technology*, Harvard University Press, 2009.

Grosfeld, Irena and Ekaterina Zhuravskaya, "Cultural vs. economic legacies of empires: Evidence from the partition of Poland," *Journal of Comparative Economics*, 2015, *43* (1), 55–75.

GUS, *Historia Polski w Liczbach - Panstwo i Spoleczenstwo*, GUS Warszawa, 2003.

_ , *Historia Polski w Liczbach - Polska w Europie*, GUS Warszawa, 2014.

Herbst, Mikolaj, "Zroznicowanie jakosci kapitalu ludzkiego w Polsce. Od czego zaleza wyniki edukacyjne?," *Studia Regionalne i Lokalne*, 2004, *17* (3), 93–104.

_ **and Steven Rivkin**, "Divergent historical experiences and inequality in academic achievement: the case of Poland," *The Journal of Socio-Economics*, 2012.

_ **, Jan Herczynski, and Anthony Levitas**, *Finansowanie oswiaty w Polsce: diagnoza, dylematy, mozliwosci*, Wydawnictwo Naukowe Scholar, 2009.

Herczynski, Jan and Aneta Sobotka, "Siec szkol podstawowych i gimnazjow w Polsce 2007-2012 - *Draft version of the report*," Technical Report, Instytut Badań Edukacyjnych 2013.

Hijmans, Robert J, Susan E Cameron, Juan L Parra, Peter G Jones, Andy Jarvis et al., "Very high resolution interpolated climate surfaces for global land areas," *International Journal of Climatology*, 2005, *25* (15), 1965–1978.

Hryniewicz, Janusz T, "Polityczne i gospodarcze nastepstwa zaboru pruskiego," *Studia Regionalne i Lokalne*, 2003, *13* (3), 61–78.

Huillery, Elise, "History matters: The long-term impact of colonial public investments in French West Africa," *American Economic Journal: Applied Economics*, 2009, pp. 176–215.

i Miquel, Gerard Padro, "The control of politicians in divided societies: The politics of fear," *The Review of Economic Studies*, 2007, *74* (4), 1259–1274.

Imbens, Guido W and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 2008, *142* (2), 615–635.

Iyer, Lakshmi, "Direct versus indirect colonial rule in India: Long-term consequences," *The Review of Economics and Statistics*, 2010, *92* (4), 693–713.

Jha, Saumitra, "Trade, institutions, and ethnic tolerance: Evidence from South Asia," *American Political Science Review*, 2013, *107* (04), 806–832.

Karaja, Elira, "The Rule of Karlowitz. Fiscal Change and Institutional Persistence," Technical Report, mimeo 2013.

Kashin, K. and D. Ziblatt, "Administrative district boundaries, Germany, 1890s," Cambridge, Mass.: Harvard Geospatial Library. Retrieved from http://nrs.harvard.edu/urn-3:HUL.gisdata:013361293. 2012.

Kassof, Grian, "EDUCATION," *Encyclopedia of Russian History*, 2004, p. 500.

Keele, Luke and Rocıo Titiunik, "Geographic boundaries as regression discontinuities," in "APSA 2011 Annual Meeting Paper" 2011.

Labuda, Gerard, *Polska granica zachodnia: tysiac lat dziejow politycznych*, Wydawn. Poznanskie, 1971.

Lamberti, Marjorie, *State, society, and the elementary school in imperial Germany*, Oxford University Press New York, 1989.

Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, June 2010, *48* (2), 281–355.

Lee, J, W Grigg, and G Dion, "The nation's report card: Mathematics 2007-National Assessment of Educational Progress at grades 4 and 8 [NCES 2007-494]," *Washington, DC: National Center for Education Statistics*, 2007.

Meissner, Andrzej and Czeslaw Majorek, *Pedagogika nowego wychowania w Polsce u schylku XIX iw pierwszej polowie XX wieku: podstawowe przejawy i wspolczesne odniesienia*, Wydawictwo Wyzszej Szkoly Pedagogicznej, 2000.

Michalopoulos, Stelios and Elias Papaioannou, "Pre-Colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 2013, *81* (1), 113–152.

MPIDR and CGG, "MPIDR Population History GIS," 2012.

Patacchini, Eleonora and Yves Zenou, "Neighborhood Effects and Parental Involvement in the Intergenerational Transmission of Education," *Journal of Regional Science*, 2011, *51* (5), 987–1013.

Porta, Rafael La, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny, "The quality of government," *Journal of Law, Economics, and Organization*, 1999, *15* (1), 222–279.

Putnam, Robert D, Robert Leonardi, and Raffaella Y Nanetti, *Making democracy work: Civic traditions in modern Italy*, Princeton University Press, 1994.

Sakalli, Seyhun Orcan, "Coexistence, Polarization, and Development: The Armenian Legacy in Modern Turkey," 2014.

Snyder, Timothy, *"Ukrainians and Poles" in: The Cambridge History of Russia: Volume 2, Imperial Russia, 1689-1917*, Vol. 2, Cambridge University Press, 2006.

Spolaore, Enrico and Romain Wacziarg, "Fertility and modernity," *Discussion Papers Series*, 2014, (779).

Steele, Claude M and Joshua Aronson, "Stereotype threat and the intellectual test performance of African Americans.," *Journal of Personality and Social Psychology*, 1995, *69* (5), 797.

Sulimierski, Filip, Bronisław Chlebowski, and Władysław Walewski, *Słownik geograficzny Królestwa Polskiego i innych krajów słowiańskich*, Vol. 14, Filipa Sulimierskiego i Władysława Walewskiego, 1895.

Wandycz, Piotr Stefan, *The lands of partitioned Poland, 1795-1918*, Vol. 7, University of Washington Press, 1974.

Whittaker, Cynthia H, *The origins of modern Russian education: an intellectual biography of Count Sergei Uvarov, 1786-1855*, Northern Illinois University Press, 1984.

Wysokinska, Agnieszka, "Importance of History for Economic Development Lessons from Natural Experiment of History." PhD dissertation, European University Institute 2011.

Zubrzycki, Jerzy, "Emigration from Poland in the nineteenth and twentieth centuries," *Population Studies*, 1953, *6* (3), 248–272.

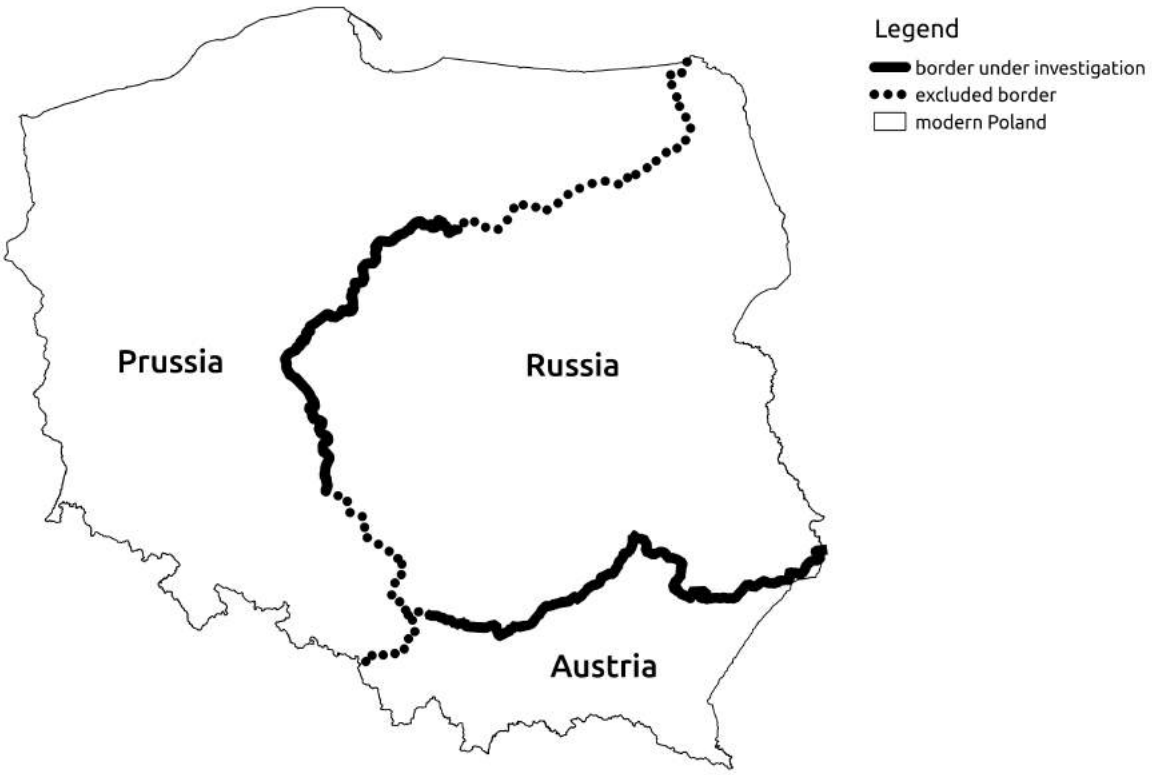

Figure 1: The Partitions of Poland 1815-1918

Note: the map shows the borders of the 19th century Partitions of Poland layered on the current map of Poland. The border under investigation are marked by the solid line, the excluded area by the dashed line. Silesia and Eastern Prussia were excluded because they belonged to Germany in the inter war period and were destination points for the massive post-WWII migration. Source: own visualization based on GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

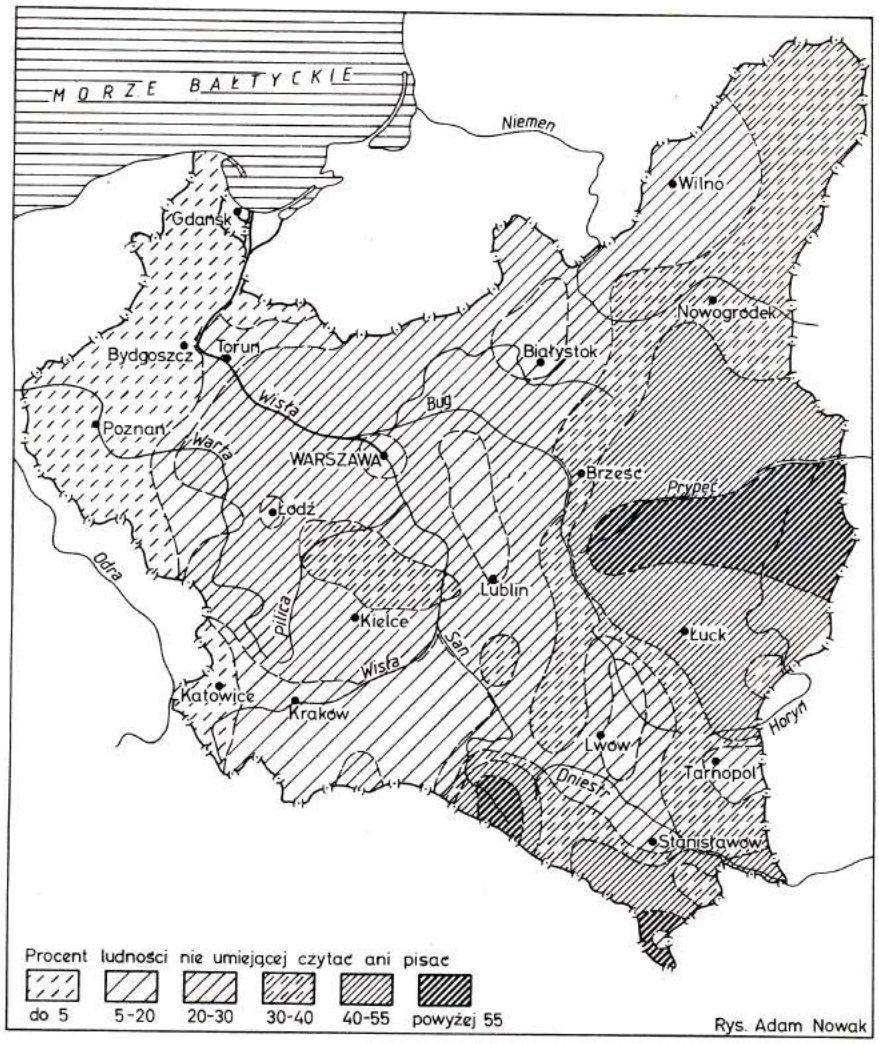

Figure 2: Illiteracy levels in the interwar Poland (1931)

Note: the map shows the share of people who cannot read and write in the interwar Poland (1931). The legend at the bottom describes the illiteracy levels. "do 5" means less than 5% and "powyżej 55" means more than 55%. Source: an illustration from Henryk Zieliński, "Historia Polski 1914-1939", Wydawnictwo Ossolineum, 1983 via http://pl.wikipedia.org/wiki/Analfabetyzm.

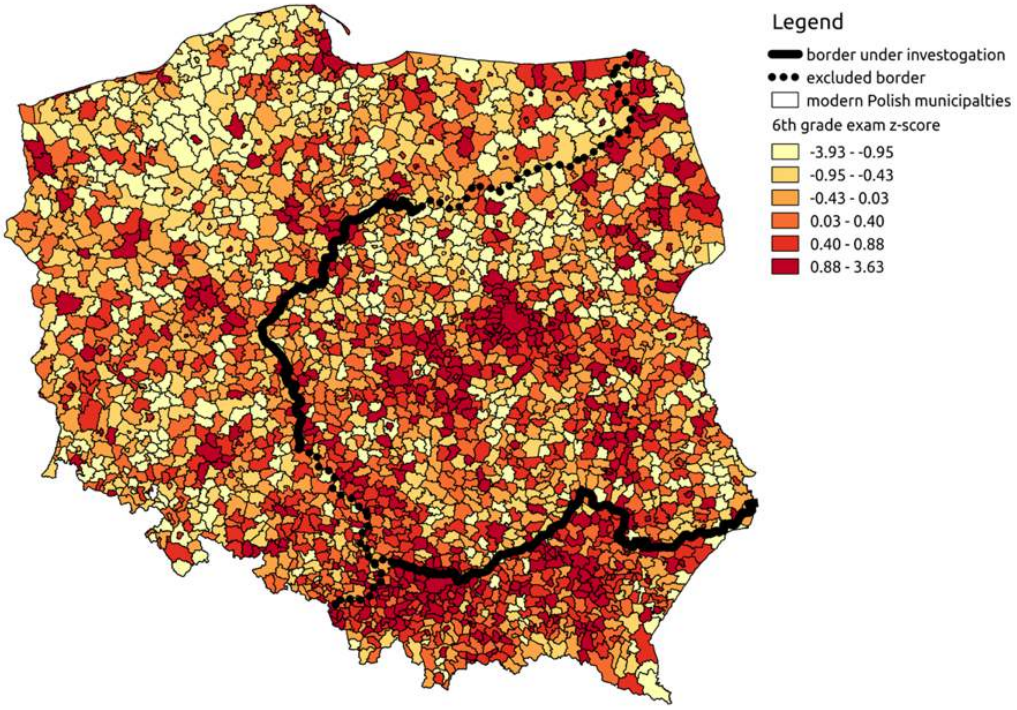

Figure 3: The distribution of the 6-th grade exam scores in 2011

Note: the map shows the distribution of the 6th grade low stake exam score in 2011 at the municipality level. The borders of the 19th century Partitions of Poland layered on the current map of Poland. The border under investigation are marked by the solid line, the excluded area by the dashed line. Silesia and Eastern Prussia were excluded because they belonged to Germany in the inter war period and were destination points for the massive post-WWII migration. Source: own visualization based on the Central Board of Examination data and GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

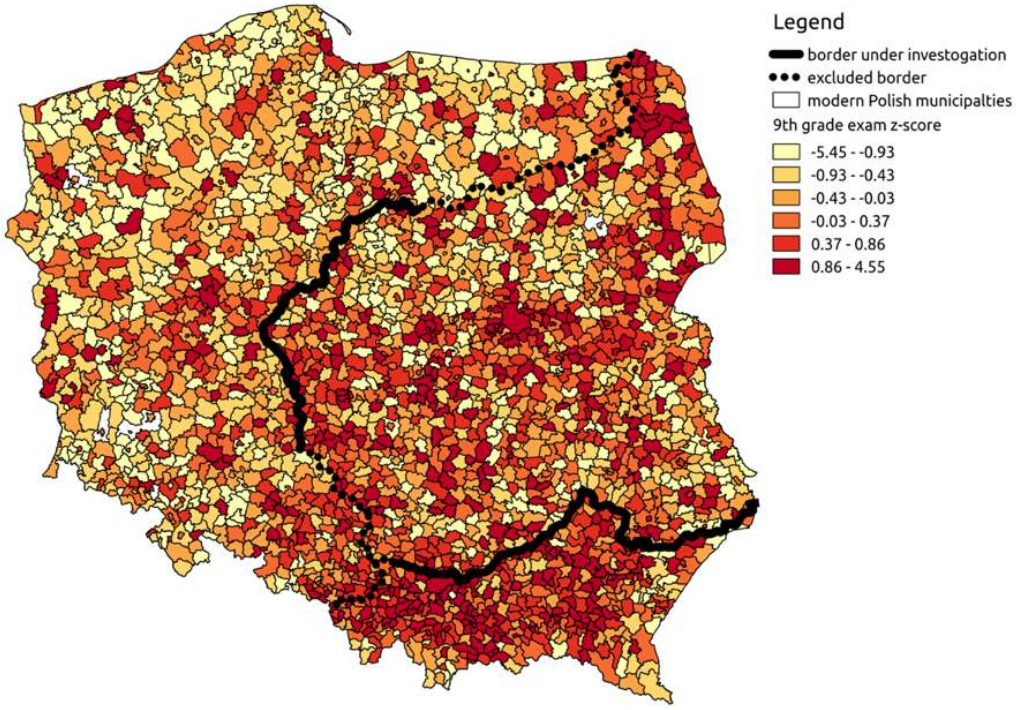

Figure 4: The distribution of the 9-th grade exam scores (math and science) in 2011

Note: the map shows the distribution of the 9th grade high stake exam score in math and science in 2011 at the municipality level. The borders of the 19th century Partitions of Poland layered on the current map of Poland. The border under investigation are marked by the solid line, the excluded area by the dashed line. Silesia and Eastern Prussia were excluded because they belonged to Germany in the inter war period and were destination points for the massive post-WWII migration. Source: own visualization based on the Central Board of Examination data and GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

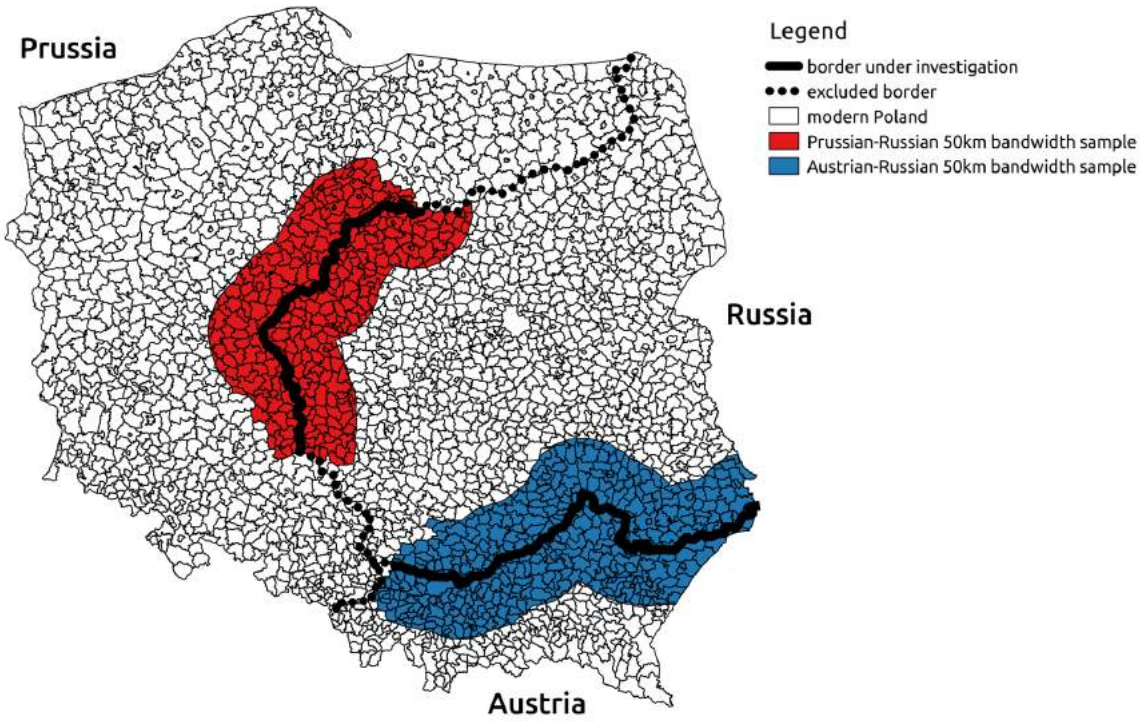

Figure 5: The total baseline sample

Note: the map shows the sample under investigation, based on the 50km bandwidth around the former borders of the Partitions of Poland, including the rural and urban areas. The borders of the 19th century Partitions of Poland layered on the current map of Poland. The border under investigation are marked by the solid line, the excluded area by the dashed line. Silesia and Eastern Prussia were excluded because they belonged to Germany in the inter war period and were destination points for the massive post-WWII migration. Source: own visualization based on GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

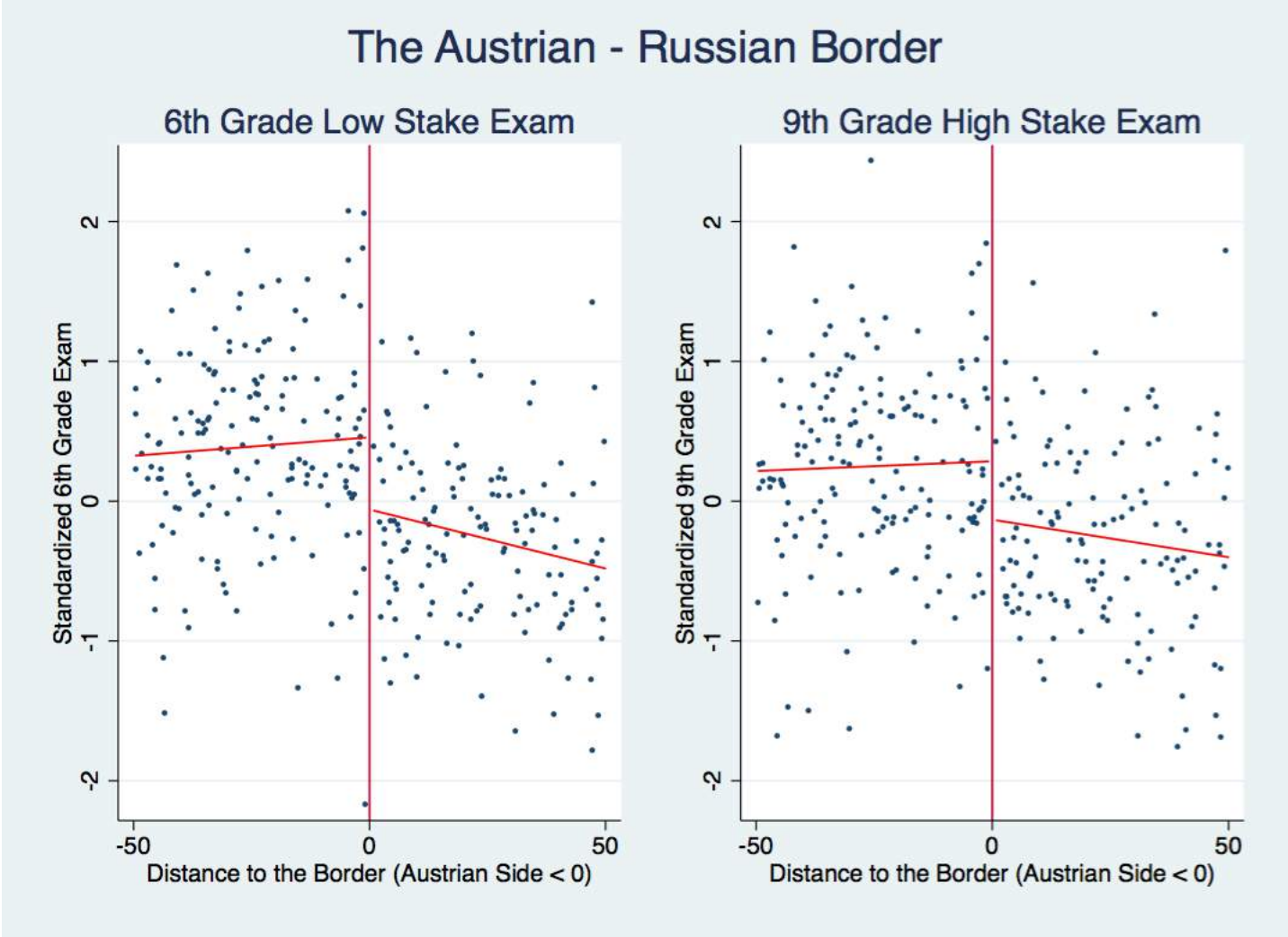

Figure 6: The one-dimensional regression discontinuity: Austrian-Russian border.

Note: the left panel of the graph shows the municipality averages of standardized 6th grade exam score from 2005-2011 plotted against distance to the Austrian-Russian border. The right panel shows analogous graph for the municipality averages of standardized 9th grade exam score 2005-2011. Negative distance is for the Austrian side. The graph uses the rural municipalities only.

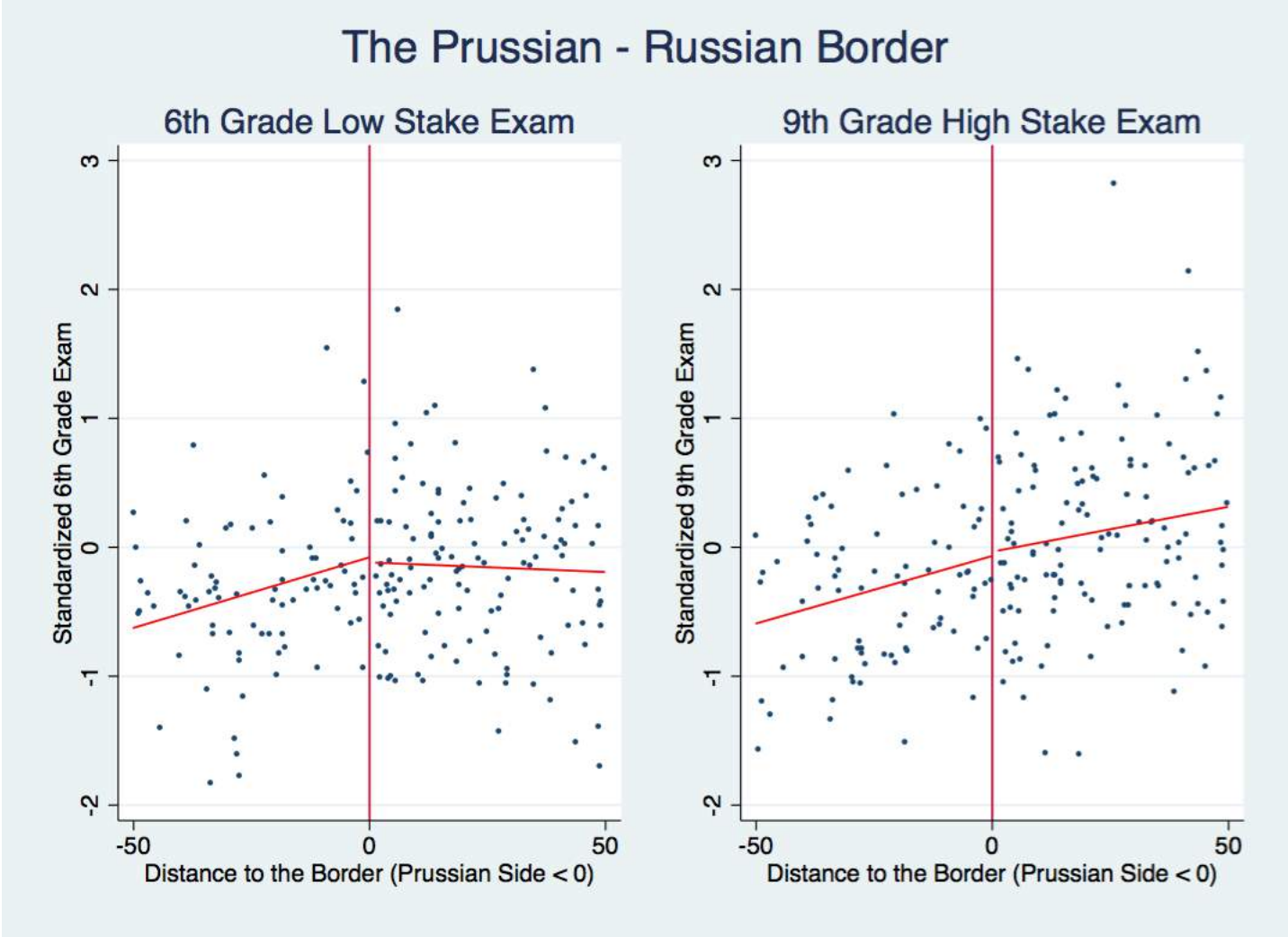

Figure 7: The one-dimensional regression discontinuity: Prussian-Russian border

Note: the left panel of the graph shows the municipality averages of standardized 6th grade exam score from 2005-2011 plotted against distance to the Prussian-Russian border. The right panel shows analogous graph for the municipality averages of standardized 9th grade exam score 2005-2011. Negative distance is for the Prussian side. The graph uses the rural municipalities only.

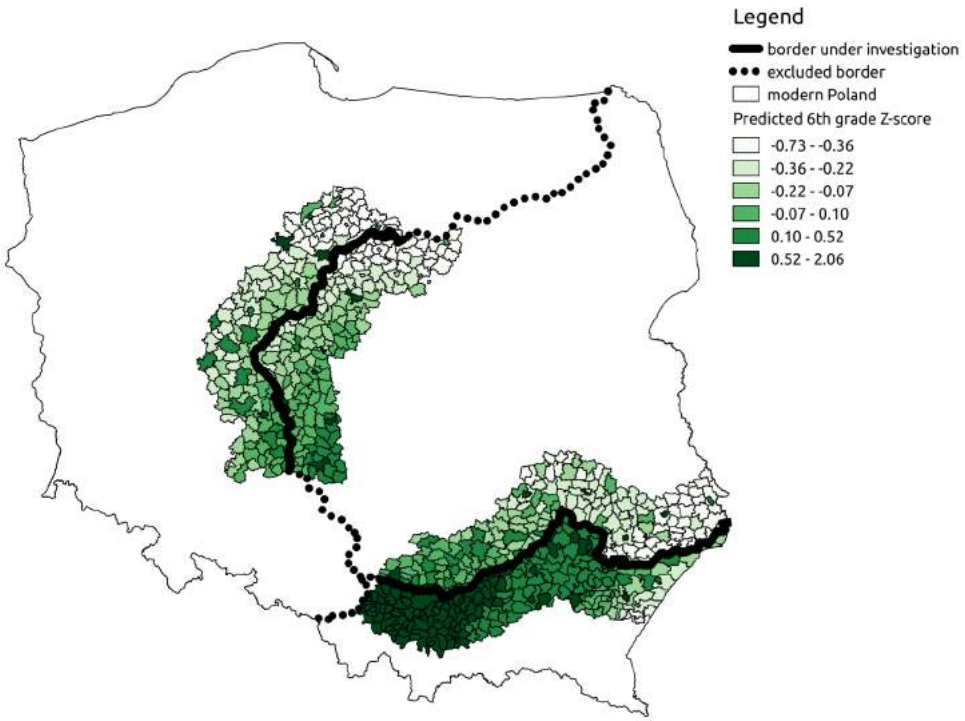

Figure 8: Predicted levels of the 6th grade exam Z-score for 2011

Note: the map shows the predicted values of the 6th grade exam Z-score for 2011, based on the regressions specified by Equation (1) for the whole sample, with the quadratic polynomial of longitude and latitude, the partition dummy, the geographic covariates and the set of population size dummies. The borders of the 19th century Partitions of Poland layered on the current map of Poland. The border under investigation are marked by the solid line, the excluded area by the dashed line. Silesia and Eastern Prussia were excluded because they belonged to Germany in the inter war period and were destination points for the massive post-WWII migration. Source: own visualization based on GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

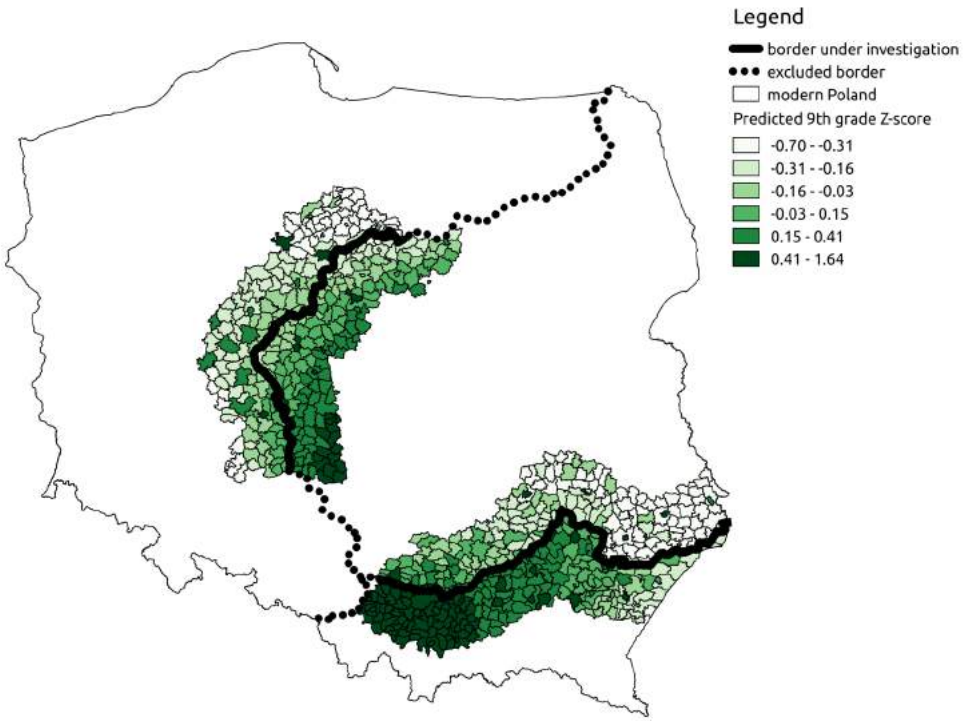

Figure 9: Predicted levels of the 9th grade exam Z-score for 2011

Note: the map shows the predicted values of the 9th grade exam Z-score for 2011, based on the regressions specified by Equation (1) for the whole sample, with the quadratic polynomial of longitude and latitude, the partition dummy, the geographic covariates and the set of population size dummies. The borders of the 19th century Partitions of Poland layered on the current map of Poland. The border under investigation are marked by the solid line, the excluded area by the dashed line. Silesia and Eastern Prussia were excluded because they belonged to Germany in the inter war period and were destination points for the massive post-WWII migration. Source: own visualization based on GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

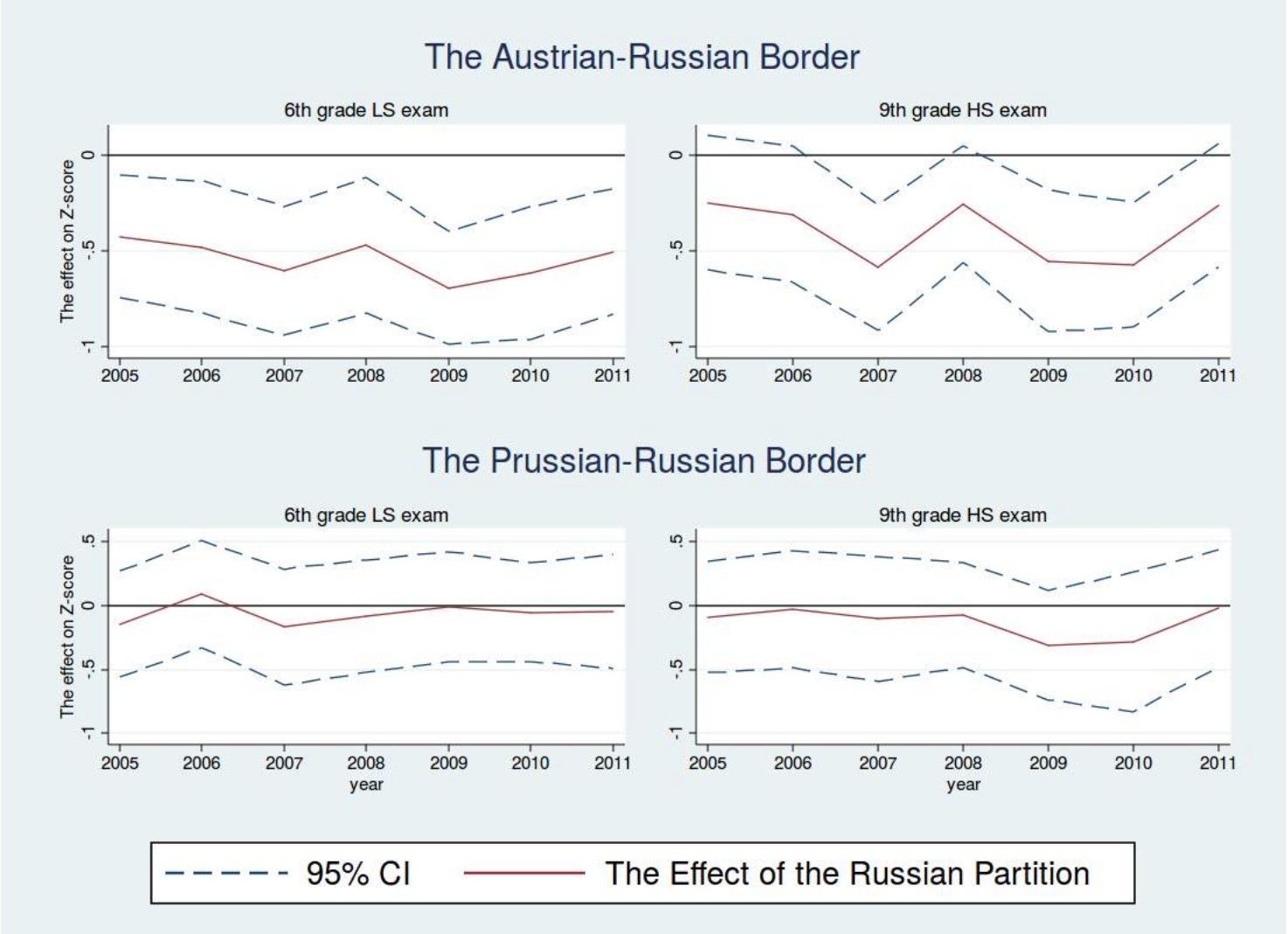

Figure 10: Estimations year-by-Y=year

Note: the figures present results of the OLS estimations of Equation (1) or each year separately with the quadratic polynomial of longitude and latitude, the rural sample and the geographic control variables. Only the coefficient and confidence intervals of the Partition dummy (variable of interest) is presented. It takes value one for the former Russian areas and zero for either the former Austrian or Prussian areas.

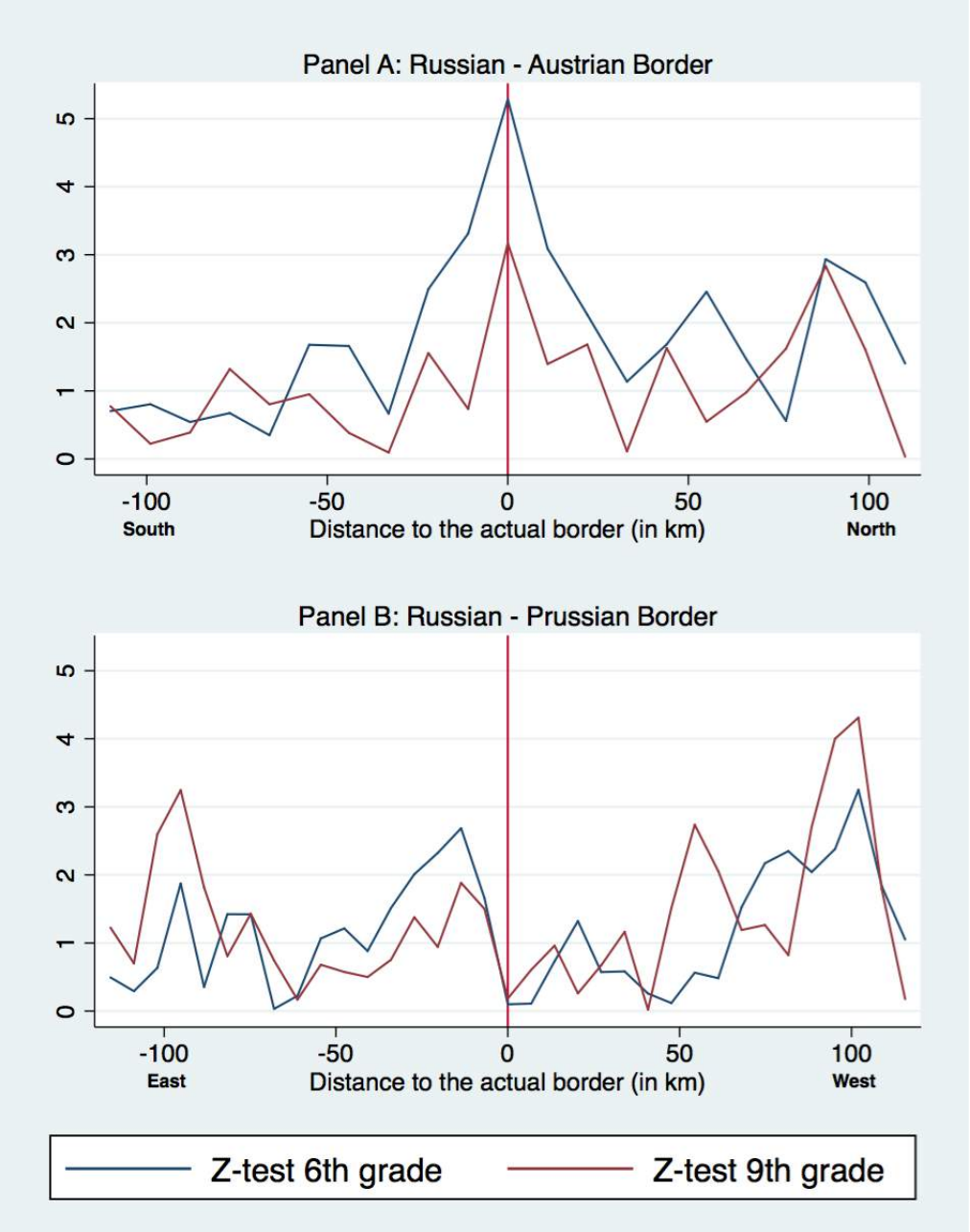

Figure 11: The placebo experiments

In Panel A I artificially move step-by-step the Austrian-Russian border by 5km to the North and to the South (negative distance). In Panel B I move step-by-step the Prussian-Russian border by 5km to the West and to the East (negative distance). For each placebo border I calculate the Z-test (ratio of a coefficient and a corresponding robust standard errors) for the placebo Partition dummy coefficient from the baseline specification of the regression specified as in Equation (1).

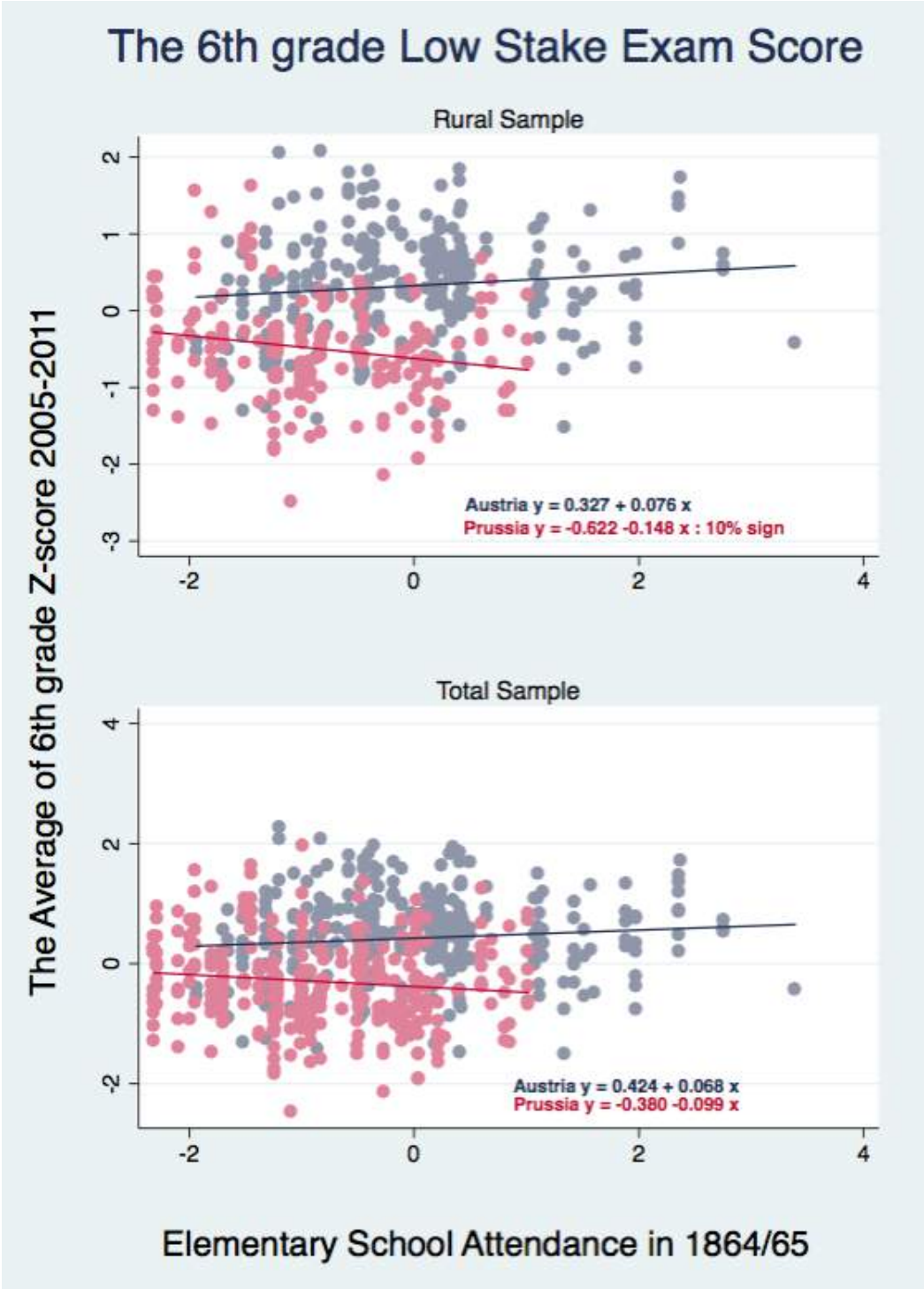

Figure 12: Historical educational attainment and modern performance of students

Note: the figures present the standardized elementary school educational attainment in 1864/65 (x-axis) results plotted against the municipality average (2005-2011) of the standardized 6th grade low stake exam (y-axis). The former Austrian partition is denoted by grey color, the former Prussian by red color. The line is a fitted line from a regression of the 6th grade exam score on the educational attainment. The sample excludes territories which were not part of Poland between 1918-1945. The top panel shows the modern rural areas only, the bottom panel the total sample.

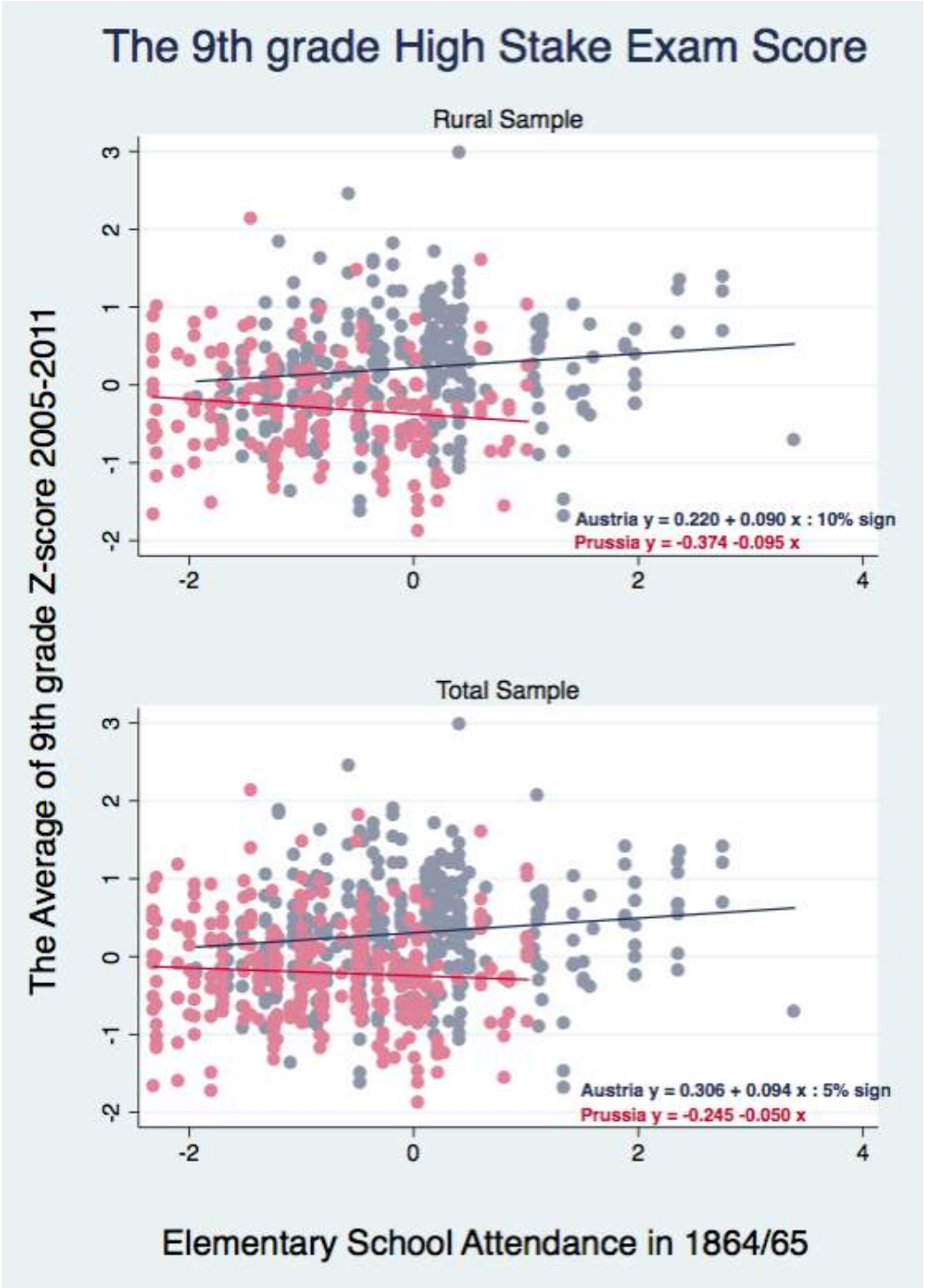

Figure 13: Historical educational attainment and modern performance of students

Note: the figures present the standardized elementary school educational attainment in 1864/65 (x-axis) results plotted against the municipality average (2005-2011) of the standardized 9th grade high stake exam (y-axis). The former Austrian partition is denoted by grey color, the former Prussian by red color. The line is a fitted line from a regression of the 9th grade exam score on the educational attainment. The sample excludes territories which were not part of Poland between 1918-1945. The top panel shows the modern rural areas only, the bottom panel the total sample.

Table 1: Historical Characteristics of the Partitions

Characteristic / Partition:	Russian	Austrian	Prussian
General Characteristics			
Agriculture	Advanced	Least Advanced	Most Advanced
Industry	Most Advanced	Least Advanced	Advanced
Law Enforcement	Lowest	Normal	Highest
Organization of the Poles	Low	High	High
Quality of Bureaucracy	Low	High	High
Educational System			
Origin	None	Prussian	Prussian
Introduction	N/A	mid 19th	early 19th
Length of compulsory education	None	8 years	8 years
School structure	Various	4+4+	8+
Financing	Various	Local	Local
Language	Russian	Polish	German
Curriculum	Russian	Polish	German
Universities	Russian	Polish	None
Ethnic Policy	Russification	Tolerance	Germanization

Table 2: Comparison of The 19th Century Educational Outcomes

Partition / Year :	1840'	1850'	1860'	1870'	1880'	1890'	1900'	1910'
Elementary School Enrollment								
Russian	-	-	-	-	-	-	18%	25%
Austrian	-	-	-	-	67%	77%	83%	86%
Prussian	62%	-	94%	-	-	-	-	-
Elementary School Students as % of Total Population								
Russian	1.3%	1.4%	2.3%	2.3%	1.9%	2.4%	2.9%	3.6%
Austrian	1.6%	1.8%	3%	3.6%	6.9%	9.7 %	11.4%	13.5%
Prussian	12.1%	-	14.3%	-	16.6%	17.4%	19%	19.3%
Total Area per Elementary School in km^2								
Russian	-	-	-	-	-	-	-	26.9
Austrian	-	-	-	-	-	-	-	12.8
Prussian	-	-	-	-	-	-	-	9.8
Elementary School Teachers per 1000 Population Aged 5-15								
Russian	-	-	-	-	-	-	-	2
Austrian	-	-	-	-	-	-	-	11
Prussian	-	-	-	-	-	-	-	13
Elementary School Pupils per Teacher								
Russian	55	-	55	-	-	49	54	56
Austrian	-	42	72	-	-	104	87	79
Prussian	-	-	-	-	91	82	73	70
Share of Population 9< who can Read								
Russian	-	-	18%	-	-	41%	-	-
Austrian	-	-	-	-	-	-	-	69%[a]
Prussian	-	-	-	-	-	-	-	95%[b]

Notes: *a*: share of population 11<. Excludes territories from modern Ukraine. *b* data only for Śląsk Cieszyński.; Otherwise, Austrian is the whole Galicia; Prussian is the Grand Duchy of Poznań; Russian is the Congress Kingdom. Source: GUS (2003) and GUS (2014).

Table 3: Descriptive Statistics

Variable / Partition:	Russian-Prussian					Russian-Austrian				
	Prussian		Russia		Diff	Austrian		Russian		Diff
	mean (1)	sd (2)	mean (3)	sd (4)	(1)-(3)	mean (6)	sd (7)	mean (8)	sd (9)	(6)-(8)
6th grade exam (2011)	24.1	1.46	24	1.54	.146	25.3	1.43	24.2	1.6	1.18***
9th grade exam (2011)	22.1	1.91	23.1	2.2	-1***	23.9	1.88	22.7	2.14	1.19***
Higher Education (2002 in %)	4.17	1.3	3.7	.926	.47**	4.89	1.88	4.16	1.31	.73***
Kindergarten attendance (2011 in %)	58.6	14.7	58.2	13.4	.45	60.3	13	54.7	13.6	5.6***
Sec. School Scholarization (2011 in %)	89.8	15.5	93.2	17.2	-3.48	91.4	8.41	93.7	10.2	-2.36*
Additional Lessons (2009)	26.3	11.1	23.5	12	2.87+	23.6	10.8	23.1	11.3	.56
Class size (2009)	16.2	2.6	14.6	2.7	1.6***	15.3	2.4	13.8	2.62	1.51***
Expenditures per capita (2011 in PLN)	3172	789	3111	617	60.5	3204	735	3079	685	125
Edu. Expenditures per capita (2011 in PLN)	1105	173	1147	188	-42.3	1163	191	1073	220	89.6***
Population (2011)	7277	3313	6205	2533	1071**	11594	7935	6379	3143	5214***
Population density (2011)	59.2	22.6	57.2	18.3	2	136	81	61.1	28.8	75***
Migration Balance (2011)	1.96	7.65	-0.18	6.05	2.14*	1.78	5.71	-2.11	3.87	3.88***
People aged 0-18 (2011 in %)	21.5	1.29	20.7	1.42	.85 ***	21.2	2.02	18.6	1.56	2.63***
Unemployment(2011 in %)	9.11	.336	10.2	.376	-1.09*	9.28	.311	9.39	.297	-.11
Agriculture (2010 in %)	10.79	8.66	4.41	6.09	6.38***	3.73	4.7	4	5.38	-.27
Altitude (in meters)	106	36.9	121	31.7	-14.4**	265	101	237	54.3	28.4**
Precipitation (in mm)	541	18.3	540	21.5	1.38	672	75.5	603	51.5	69***
Temperature (in C°)	7.93	0.34	8.07	0.44	-.14*	7.83	0.5	7.57	0.34	.26***
Number of municipalities	80		126			164		137		

Notes: Means and standard deviations for the sample of rural municipalities, located at most 50km either from the former Russian-Prussian or Russian-Austrian border. Municipalities located in Śląskie, Warmińsko-Mazurskie and Opolskie vovoidships are excluded. Definitions of the variables are given in Table A1.

Table 4: Geographic Differences

Dep. Variable:	Altitude (m)		Precipitation (mm)		Temperature (C°)	
	(1)	(2)	(3)	(4)	(5)	(6)
Russian - Austrian Border						
Panel A : Quadratic Polynomial in Latitude and Longitude						
Partition Effect	79	82	30.8	33.3	-.43	-.47
(Russia=1)	(11.5)***	(10)***	(4.2)***	(3.7)***	(.07)***	(.06)***
R^2	.58	.58	.93	.93	.30	.31
Municipalities	301	373	301	373	301	373
Panel B: Quadratic Polynomial in Distance						
Partition Effect	8.9	6.8	8.9	6.2	.03	.05
(Russia=1)	(20.1)	(18.7)	(17.4)	(16.3)	(.12)	(.11)
R^2	.30	.30	.37	.36	.34	.35
Municipalities	301	373	301	373	301	373
Russian - Prussian Border						
Panel C : Quadratic Polynomial in Latitude and Longitude						
Partition Effect	-4	-10.6	-1.9	-3.5	0.4	0.9
(Russia=1)	(3.3)	(3.03)***	(2.2)	(1.8)*	(.02)*	(.02)***
R^2	.83	.81	.85	.84	.95	.94
Municipalities	206	302	206	302	206	302
Panel D : Quadratic Polynomial in Distance						
Partition Effect	-5.6	-6.7	-3.5	-8.4	-0.01	0.7
(Russia=1)	(10.8)	(9.8)	(8.4)	(7)	(.14)	(.11)
R^2	.09	.07	.02	.03	.05	.03
Municipalities	206	302	206	302	206	302
Sample	rural	all	rural	all	rural	all

Notes: Robust and clustered at the municipality level standard errors are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level and * at the 5%. Columns (1) to (2) - the dependent variable is the average altitude in meters; columns (3) to (4) the average annual precipitation in millimeters; columns (5) to (6) the average annual temperature in Celsius degrees. Table presents estimates of the coefficient β from the regression (1)) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian (Panel A, C) or Prussian (Panel B, D) territories. In addition the regressions include a quadratic polynomial in latitude and longitude (Panel A, B) or a quadratic polynomial in distance (Panel C, D). All the regressions use 50 km bandwidth.

Table 5: Baseline Regressions

Dep. Variable:	6th grade LS exam				9th grade HS exam			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Russian - Austrian Border								
Panel A : Quadratic Polynomial in Latitude and Longitude								
Partitions Effect	-.615	-.542	-.592	-.536	-.422	-.396	-.393	-.392
(Russia=1)	(.115)***	(.121)***	(.112)***	(.112)***	(.113)***	(.128)**	(.101)***	(.12)***
R^2	.26	.3	.28	.32	.19	.2	.21	.22
Municipalities	301	301	373	373	301	301	373	373
Mun. X Time	2107	2107	2606	2606	2106	2106	2605	2600
Panel B: Quadratic Polynomial in Distance								
Partitions Effect	-.434	-.466	-.472	-.481	-.225	-.266	-.333	-.347
(Russia=1)	(.248)+	(.228)*	(.226)*	(.208)*	(.223)	(.211)	(.225)	(.212)
R^2	.14	.19	.2	.24	.09	.16	.15	.2
Municipalities	301	301	373	373	301	301	373	373
Mun. X Time	2107	2107	2606	2606	2106	2106	2605	2605
Russian - Prussian Border								
Panel C : Quadratic Polynomial in Latitude and Longitude								
Partitions Effect	-.034	-.057	-.129	-.147	-.07	-.129	-.041	-.088
(Russia=1)	(.151)	(.151)	(.121)	(.124)	(.168)	(.165)	(.128)	(.129)
R^2	.06	.07	.11	.11	.11	.13	.11	.14
Municipalities	206	206	302	302	206	206	302	302
Mun. X Time	1442	1442	2114	2114	1442	1442	2114	2114
Panel D : Quadratic Polynomial in Distance								
Partitions Effect	-.392	-.371	-.322	-.323	-.14	-.115	-.132	-.123
(Russia=1)	(.229)+	(.232)	(.197)	(.196)	(.235)	(.243)	(.183)	(.19)
R^2	.04	.08	.08	.12	.07	.09	.07	.1
Municipalities	206	206	302	302	206	206	302	302
Mun. X Time	1442	1442	2114	2114	1442	1442	2114	2114
Geo. Controls	no	yes	no	yes	no	yes	no	yes
Sample	rural	rural	all	all	rural	rural	all	all

Notes: Robust and clustered at the municipality level standard errors are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level, * at the 5% and + at the 10%. Columns (1) to (3) - the dependent variables are the 6th grade low-stake exam score; Columns (4) to (6)the mathematics and science 9th grade high-stake exam score. Table presents estimates of the coefficient β from the regression (1) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian (Panel A, C) or Prussian (Panel B, D) territories. In addition the regressions include a quadratic polynomial in latitude and longitude (Panel A, B) or a quadratic polynomial in distance (Panel C, D) and a set of geographic covariates (columns 2,4, 6 and 8). All the regressions use 50 km bandwidth.

Table 6: Discontinuities with log of covariates as dependent variables.

Dep. Variable / Border:	Prussian-Russian (1)	Austrian-Russian (2)
Panel A : Time-Variant variables		
Expenditures	-.028	-.032
	(.036)	(.025)
Educational Expenditures	.035	-.010
	(.036)	(.025)
Unemployment Rate	.081	-.265
	(.071)	(.062)***
Sec. School Scholarization	.056	.036
	(.040)	(.020)*
Population	-.154	-.366
	(.108)	(.095)***
Population 0-18	.011	-.054
	(.015)	(.015)***
Population Density	.143	-.255
	(.092)	(.082)***
Municipalities X Time	1442	2105
Municipalities	206	301
Panel B : Time-Invariant variables		
Agriculture	-.875	.110
	(.304)***	(.262)
Higher Education	-.042	-.154
	(.069)	(.060)***
Additional Classes	.057	.119
	(.110)	(.103)
Class Size	-.111	-.037
	(.039)***	(.038)
Educational Value Added	.074	.004
	(.419)	(.298)
Municipalities	206	298
Geographic Controls	yes	yes
Sample	rural	rural

Notes: Robust and clustered at the municipality level standard errors are reported in the parentheses. *** denotes significance at the 0.1% level, ** at the 1% level and * at the 5%. Table presents estimates of the coefficient β from the regression (specified as (1)) of *logarithms* of various dependent variables (except educational value added) on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian (column 2) or Prussian (column 1). Column (1) shows the effect of the Rusian Empire for the Prussian-Russian border, Column (2) for the Austrian-Russian. In addition the regressions include a quadratic polynomial in latitude and longitude and geographic covariates. The dependent variables are explained in Table A1. All the regressions use 50 km bandwidth.

Table 7: Proxies for Social Norms: Social Diagnosis Survey. Marginal Effects at the Border.

Dep. Variable:	Education - Important for a Good Life (1=Yes)		Satisfied with Received Education (1=Yes)		University as a Desired Degree for a Child (1=Yes)		Agree that Corporal Punishment is Important for a Child Dev. (1=Yes)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Russian - Austrian Border								
Partitions Effect (Russia=1)	-.047 (.022)*	-.054 (.021)**	-.031 (.046)	-.065 (.044)	.176 (.083)*	.22 (.072)**	-.053 (.008)***	-.133 (.04)***
R^2	.16	.16	.04	.04	.06	.07	.02	.02
Observations	8364	8364	8259	8259	2496	2496	8351	8351
Russian - Prussian Border								
Partitions Effect (Russia=1)	.038 (.016)*	.033 (.017)*	-.052 (.059)	-.074 (.057)	.038 (.086)	.083 (.087)	.128 (.035)***	.102 (.038)**
R^2	.15	.15	.02	.02	.08	.091	.01	.02
Observations	7853	7853	7763	7763	2149	2149	7884	7884
Background Controls	yes	yes	yes	yes	yes	yes	yes	yes
Geographic Controls	yes	yes	yes	yes	yes	yes	yes	yes
Modern Controls	no	yes	no	yes	no	yes	no	yes
Sample	all	all	all	all	all	all	all	all
Estimator	Probit	Probit	Probit	Probit	Probit	Probit	OLogit	OLogit

Notes: Robust and clustered standard errors (at the county level) are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level, * at the 5% and + at the 10%. Dependent variables are described in Table A3. Table presents estimates of the marginal effects at the border of the coefficient β from the regression (1) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian or Prussian territories. In addition the regressions include a quadratic polynomial of distance and a set of exogenous and endogenous covariates. The endogenous control variables are listed and explained in Table A1. All the regressions use 100 km bandwidth.

Table 8: Proxies for Social Norms: LiTS and EVA. The Average Marginal Effects

Dep. Variable:	First or Second Priority of Governmental Spending on Public Education (1=Yes)		Intelligence and Skills Important in Life Success (1=Yes)		Log Spending on Educations		Family Tradition Important in School Selection (1=Yes)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Russian - Austrian Border								
Partitions Effect	-.024	-.19	.116	.203	.177	.158	-.055	-.07
(Russia=1)	(.069)	(.088)*	(.058)*	(.067)**	(.23)	(.343)	(.032)+	(.053)
R^2	.06	.06	.04	.09	.16	.25	.01	.03
Observations	602	602	602	587	233	233	802	802
Russian - Prussian Border								
Partitions Effect	.076	-.048	-.07	-.174	.353	-.448	.027	-.022
(Russia=1)	(.049)	(.076)	(.074)	(.071)*	(.216)	(.74)	(.029)	(.037)
R^2	.08	.1	.03	.1	.3	.37	.01	.05
Observations	461	461	461	461	166	166	1050	1050
Background Controls	yes	yes	yes	yes	yes	yes	yes	yes
Geographic Controls	yes	yes	yes	yes	yes	yes	yes	yes
Modern Controls	no	yes	no	yes	no	yes	no	yes
Sample	all	all	all	all	all	all	all	all
Estimator	Probit	Probit	Probit	Probit	OLS	OLS	OLogit	OLogit
Source	LiTS	LiTS	LiTS	LiTS	LiTS	LiTS	EVA	EVA

Notes: Robust and clustered standard errors (at the county level) are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level , * at the 5% and + at the 10%. Dependent variables are described in Table A2. Table presents estimates of the average marginal effects of the coefficient β from the regression (1) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian or Prussian territories. Because the sample are several randomly selected PSU, the regressions does not include any polynomial of location or distance. The endogenous control variables are listed and explained in Table A1. All the regressions use 100 km bandwidth.

Table 9: Kindergarten Attendance

Dep. Variable:	Kindergarten Attendance					
	(1)	(2)	(3)	(4)	(5)	(6)
Russian - Austrian Border						
Partitions Effect	-7.16	-5.07	-3.29	-5.74	-3.37	-5.4
(Russia=1)	(1.97)***	(2.17)*	(2.09)	(1.78)***	(1.99)+	(1.95)***
R^2	.21	.21	.4	.34	.35	.5
Municipalities	301	301	301	373	373	373
Mun. X Time	2107	2107	2101	2611	2611	2606
Russian - Prussian Border						
Partitions Effect	-.227	-.071	-1.27	-.314	-.043	-.44
(Russia=1)	(2.55)	(2.57)	(2.56)	(2.22)	(2.24)	(2.04)
R^2	.36	.36	.54	.41	.41	.55
Municipalities	206	206	206	302	302	302
Mun. X Time	1442	1442	1442	2114	2114	2114
Geo. Controls	no	yes	yes	no	yes	yes
Modern Controls	no	no	yes	no	no	yes
Sample	rural	rural	rural	all	all	all

Notes: Robust and clustered at the municipality level standard errors are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level, * at the 5% and + at the 10%. The dependent variable is kindergarten attendance defined as pre-elementary schools' attendance divided by number of children aged 3-5. Table presents estimates of the coefficient β from the regression (1) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian (Panel A) or Prussian (Panel B) territories. The regressions include a quadratic polynomial in latitude and longitude, a set of geographic covariates (columns (2) and (6)) and a set of modern covariates (columns (3) and (7)). All the regressions use 50 km bandwidth.

Table 10: The Historical Data on Migration and Demographic Characteristics

Year / Partition:	Russian	Austrian	Prussian
Panel A: Migration Balance (in thousands)			
1881-1890	N/A	-74	-233
1891-1900	N/A	-169	-219
1901-1910	N/A	-224	-180
1881-1910	N/A	-468	-632
Panel B: Net Migration in 1871-1910s as % of 1910 Population			
1871-1910	11%	13%	20%
Panel C: Average of annual rate of population growth			
1846-1870	0.9%	0.5%	0.6%
1870-1897	1.6%	0.9%	0.6%
1897-1911	1.7%	1%	1%
Panel D: Share of age group in 1900			
<19	49.2%	48.7%	N/A
20-39	30.3%	28.7%	N/A
40-59	14.3%	16.7%	N/A
60<	6.2%	5.4%	N/A

Notes: Panels A and B: Austrian is Western Galicia; Prussian is the Duchy of Poznań. Panels C and D: Austrian is the whole Galicia; Prussian is the Duchy of Poznań; Russian is the Congress Kingdom. Source: GUS (2003) and Zubrzycki (1953).

Table 11: Correlations between the 19th Century Educational Attainment and Modern Performance of Students

Dep. Variable:	6th LS exam				9th HS exam			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Austria	2.23	1.65	1.88	1.75	1.77	1.67	1.53	1.64
	(.404)***	(.409)***	(.373)***	(.332)***	(.389)***	(.402)***	(.346)***	(.342)***
Attainment	-.259	-.136	-.234	-.148	-.143	-.077	-.101	-.039
	(.106)*	(.075)+	(.09)**	(.061)*	(.086)+	(.078)	(.086)	(.076)
Austria X Attainment	.282	.17	.24	.162	.172	.11	.128	.069
	(.114)*	(.085)*	(.097)*	(.07)*	(.097)+	(.089)	(.095)	(.076)
R^2	.32	.38	.25	.36	.2	.22	.17	.23
Municipalities X Time	3324	3318	4585	4587	3325	3319	4587	4581
Municipalities	475	475	656	656	475	475	656	656
Deanery/County	102	102	112	112	102	102	112	112
Geographic Controls	yes	yes	yes	yes	yes	yes	yes	yes
Modern Controls	no	yes	no	yes	no	yes	no	yes
Sample	rural	all	rural	all	rural	all	rural	all

Notes: Robust and clustered at the historical county (deanery) level standard errors are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level, * at the 5% and + at the 10%. Columns (1) to (4) - the dependent variables are the 6th grade low-stake exam score; Columns (5) to (8) the mathematics and science 9th grade high-stake exam score. Table presents estimates of the effect of the 19th century educational attainment on the dependent variables. The regressions include geographical controls. In addition. some regressions include a set of modern time-variant socio-economic covariates. The control variables are listed and explained in Table A1. The sample excludes territories which were not part of Poland between 1918-1945.

Appendix

Table A1: Variables Description: The Regression Discontinuity Design

Variable	Description	Time
Panel A: Regression Discontinuity Design - Geographic Controls		
Altitude:	The municipality average of altitude in meters.	-
Precipitation:	The municipality average (1950-2000) annual precipitation in mm.	-
Temperature:	The municipality average (1950-2000) annual temperature in C°.	-
Panel B: Regression Discontinuity Design - Endogenous Controls		
Density:	Population density.	05-11
Expenditures:	Local government (municipality) total expenditures per capita.	05-11
Educational Expenditures:	Local government (municipality) educational expenditures per capita.	05-11
Kindergarten attendance:	Rate of student pre-elementary schools' attendance.	05-11
Migration:	Migration balance per 1000 inhabitants.	05-11
Population:	Total population.	05-11
Secondary School Scholarization:	Rate of student secondary schools' attendance.	05-11
Unemployment Rate:	Share of unemployed among the active population.	05-11
Panel C: Other Variables		
Agriculture:	Share of employed in the agriculture sector among all employed.	2010
Additional Lessons:	Average number of additional lessons per elementary school.	2009
Class size:	Average class size in elementary schools.	2009
Higher Education:	Share of people with higher education.	2002
People aged 0-18:	Share of people aged 0-18.	05-11
Educational Value Added:	The estimates of the Educational Value Added (gain between 6th and 9th grade).	2013

Notes: All the variables come from the Central Statistical Office of Poland, except the variables for 2009, which come from the System of Educational Information, for the educational value added, which comes from the Educational Value Added Team and for the geographical controls, which come form *WorldClim.org*.

Table A2: Variables Description: LiTS (2006 and 2010) and EVA (2010)

Variable	**Description**
Panel A: LiTS - Outcomes	
First or Second Priority of Governmental Spending on Public Education:	"In your opinion, which of these fields should be given first or second priority for extra government spending?" with possible answer including: education, health care, housing, pensions, assisting the poor, environment protection, public infrastructure, other (the respondent could choose only one answer). The dummy equals 1 if the respondent chose education for first or second priority and 0 otherwise.
Intelligence and Skills Important for Life Success:	"In your opinion, which of the following factors is the most important to succeed in life in our country now?" with possible answer including: Effort and Hard Work; Intelligence and Skills; By Political Connections; By Breaking the Law; Other (the respondent could choose only one answer). The dummy equals 1 if the respondent chose Intelligence and Skills and 0 otherwise.
Log Spending on Education:	"Approximately how much did your household spend on education during the past 12 months?".
Panel B: LiTS - Exogenous Controls	
Gender:	Equals 1 if the respondent is a female and 0 otherwise.
Age:	Age of the respondent in years.
Having a Child:	Equals 1 if the respondent has at least one child younger than 14 years old and 0 otherwise.
Panel C: EVA - Outcomes	
Family Tradition Important in School Selection:	If parents considered an alternative school (to the local one), the question asks to select factors and sources of information which were important for the final selection of the school. Respondents could select multiple answers, family tradition is one of the possibility. The dummy equals 1 if the respondent selected family tradition.
Panel D: EVA - Exogenous Controls	
Child Gender:	Equals 1 if the child is a female and 0 otherwise.
Respondent Gender:	Equals 1 if the respondent is a female and 0 otherwise.
Age:	Age of the respondent in years.
Parent:	Equals 1 if the respondent is a parent of the child.

Table A3: Variables Description: Social Diagnosis (2011 and 2013)

Variable	**Description**
Panel A: Social Diagnosis - Outcomes	
Education - Important for a Good Life:	"What do you think is the most important for a successful and happy life?" Respondents are asked to select at most three answers, education is one of the options. The dummy equals 1 if the respondent chose education and 0 otherwise.
Satisfied with Received Education:	"Are you satisfied from your education?" the respondents could select one answer from a six-degree scale where 1 is "Very Satisfied" and 6 "Not Satisfied at all". The dummy equals 1 if the respondent choose degree "Very Satisfied", "Satisfied" or "Somehow Satisfied" and 0 otherwise.
University as a Desired Degree for a Child:	"What is the desired level of education for your child ?" the respondents could select one answer from a five-degree scale where 1 is "Primary-vocational" and 5 "Higher Education - MA". The dummy equals 1 if the respondent choose degree "Higher Education - MA" or "Higher Education - BA" and 0 otherwise.
Disagree that Corporal Punishment is Important for a Child Development:	"Do you agree with the following statement: Without corporal punishments it is impossible to rise children properly". the respondents could select one answer from a seven-degree scale where 1 is "Definitely Yes", 4 is "Neither Yes nor No" and 7 "Definitely No". The categorical variable equals 1 if the respondent choose "Definitely Yes", "Yes" or "Rather Yes"; 2 if "Neither Yes nor No"; 3 if "Rather No", "No" or "Definitely No". The reported average marginal effects show the effect on the last category (=3).
Panel B: Social Diagnosis - Exogenous Controls	
Gender:	Equals 1 if the respondent is a female and 0 otherwise
Age:	Age of the respondent in years
Size of hometown:	A categorical variable with a six-degree scale where 1 is "Cities larger than 500 thousand" and 6 is "Villages"

Table A4: Results: Polynomials in Latitude and Longitude

Dep. Variable:	6th grade LS exam			9th grade HS exam		
Polynomial / Bandwidth:	<50km	<75km	<100km	<50km	<75km	<100km
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A : Russian - Austrian Border						
Linear	-.550	-.670	-.609	-.442	-.480	-.398
	(.112)***	(.104)***	(.099)***	(.121)***	(.109)***	(.105)***
Quadratic	-.542	-.600	-.594	-.399	-.421	-.381
	(.121)***	(.111)***	(.106)***	(.128)***	(.114)***	(.110)***
Cubic	-.529	-.556	-.532	-.382	-.397	-.324
	(.119)***	(.111)***	(.107)***	(.130)***	(.118)***	(.115)***
Quartile	-.538	-.546	-.530	-.395	-.380	-.312
	(.119)***	(.113)***	(.106)***	(.128)***	(.119)***	(.114)***
Municipalities X Time	2107	2981	3688	2106	2981	3681
Municipalities	301	426	527	301	426	527
Panel B : Russian - Prussian Border						
Linear	-.030	.159	.035	-.093	.332	.241
	(.144)	(.137)	(.122)	(.160)	(.151)**	(.130)*
Quadratic	-.057	.125	.039	-.129	.310	.239
	(.151)	(.137)	(.123)	(.165)	(.152)**	(.133)*
Cubic	-.058	.096	.043	-.132	.287	.160
	(.148)	(.137)	(.127)	(.166)	(.153)*	(.136)
Quartile	-.098	.032	.047	-.173	.248	.128
	(.148)	(.138)	(.127)	(.165)	(.154)	(.137)
Municipalities X Time	1442	2135	2898	1442	2135	2894
Municipalities	206	305	414	206	305	414
Geographic Controls	yes	yes	yes	yes	yes	yes
Socio-Economic Controls	no	no	no	no	no	no
Sample	rural	rural	rural	rural	rural	rural

Notes: Robust and clustered at the municipality level standard errors are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level and * at the 5%. Columns (1) to (3) - the dependent variable is the 6th grade low-stake exam score; Columns (4) to (6) the mathematics and science 9th grade high-stake exam score. Table presents estimates of the coefficient β from the regression (1) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian (Panel A) or Prussian (Panel B) territories. The regressions use 50 km (columns (1) and (4)), 75km (columns (2) and (5)) and 100km (columns (3) and (6)) bandwidths.

Table A5: Results: Polynomials in Latitude and Longitude, the Total Sample.

Polynomial / Bandwidth:	6th grade LS exam			9th grade HS exam		
	<50km	<75km	<100km	<50km	<75km	<100km
Polynomial /	(1)	(2)	(3)	(4)	(5)	(6)
Panel A : Russian - Austrian Border						
Linear	-.592	-.688	-.640	-.445	-.481	-.434
	(.103)***	(.094)***	(.090)***	(.108)***	(.097)***	(.093)***
Quadratic	-.535	-.596	-.612	-.392	-.420	-.422
	(.112)***	(.101)***	(.096)***	(.120)***	(.104)***	(.100)***
Cubic	-.514	-.548	-.556	-.374	-.401	-.383
	(.112)***	(.104)***	(.098)***	(.122)***	(.109)***	(.105)***
Quartile	-.527	-.536	-.552	-.390	-.381	-.367
	(.112)***	(.104)***	(.098)***	(.120)***	(.110)***	(.104)***
Municipalities X Time	2606	3640	4508	2605	3641	4502
Municipalities	373	521	645	373	521	645
Panel B : Russian - Prussian Border						
Linear	-.129	.039	-.047	-.012	.237	.170
	(.117)	(.110)	(.097)	(.120)	(.115)**	(.098)*
Quadratic	-.147	.006	-.044	-.088	.207	.178
	(.124)	(.111)	(.098)	(.129)	(.117)*	(.101)*
Cubic	-.145	-.024	-.053	-.090	.191	.128
	(.123)	(.111)	(.100)	(.129)	(.117)	(.103)
Quartile	-.184	-.088	-.044	-.125	.143	.105
	(.124)	(.112)	(.101)	(.129)	(.117)	(.103)
Municipalities X Time	2114	3094	4214	2114	3094	4210
Municipalities	302	442	602	302	442	602
Geographic Controls	yes	yes	yes	yes	yes	yes
Socio-Economic Controls	no	no	no	no	no	no
Sample	all	all	all	all	all	all

Notes: Robust and clustered at the municipality level standard errors are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level and * at the 5%. Columns (1) to (3) - the dependent variable is the 6th grade low-stake exam score; Columns (4) to (6) the mathematics and science 9th grade high-stake exam score. Table presents estimates of the coefficient β from the regression (1) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian (Panel A) or Prussian (Panel B) territories. The regressions use 50 km (columns (1) and (4)), 75km (columns (2) and (5)) and 100km (columns (3) and (6)) bandwidths. The regressions use the whole sample (urban and rural).

Table A6: Results: Polynomials in Distance

Dep. Variable:	6th grade LS exam			9th grade HS exam		
Polynomial / Bandwidth:	<50km	<75km	<100km	<50km	<75km	<100km
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A : Russian - Austrian Border						
Linear	-.555	-.648	-.557	-.457	-.466	-.318
	(.164)***	(.136)***	(.128)***	(.146)***	(.122)***	(.121)***
Quadratic	-.466	-.419	-.536	-.266	-.352	-.447
	(.228)**	(.191)**	(.171)***	(.211)	(.173)**	(.156)***
Cubic	-.486	-.408	-.291	-.560	-.218	-.202
	(.332)	(.258)	(.230)	(.293)*	(.229)	(.211)
Quartile	-.379	-.428	-.374	-.458	-.228	-.243
	(.319)	(.256)*	(.227)*	(.283)	(.228)	(.206)
Municipalities X Time	2107	2981	3688	2106	2981	3681
Municipalities	301	426	527	301	426	527
Panel B : Russian - Prussian Border						
Linear	-.066	.104	.001	.019	.422	.285
	(.157)	(.134)	(.116)	(.168)	(.146)***	(.128)**
Quadratic	-.371	-.214	.070	-.115	-.158	.269
	(.232)	(.196)	(.174)	(.243)	(.204)	(.185)
Cubic	-.432	-.476	-.345	-.317	-.336	-.337
	(.310)	(.256)*	(.226)	(.324)	(.266)	(.234)
Quartile	-.788	-.420	-.532	.063	.047	-.332
	(.444)*	(.327)	(.278)*	(.439)	(.337)	(.289)
Municipalities X Time	1442	2135	2898	1442	2135	2894
Municipalities	206	305	414	206	305	414
Geographic Controls	yes	yes	yes	yes	yes	yes
Socio-Economic Controls	no	no	no	no	no	no
Sample	rural	rural	rural	rural	rural	rural

Notes: Robust and clustered at the municipality level standard errors are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level and * at the 5%. Columns (1) to (3) - the dependent variable is the 6th grade low-stake exam score; Columns (4) to (6) the mathematics and science 9th grade high-stake exam score. Table presents estimates of the coefficient β from the regression (1) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian (Panel A) or Prussian (Panel B) territories. The regressions use 50 km (columns (1) and (4)), 75km (columns (2) and (5)) and 100km (columns (3) and (6)) bandwidths.

Table A7: Results: Polynomials in Latitude and Longitude, including Socio-Economic Covariates.

Dep. Variable:	6th grade LS exam			9th grade HS exam		
Polynomial / Bandwidth:	<50km	<75km	<100km	<50km	<75km	<100km
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A : Russian - Austrian Border						
Linear	-.356	-.440	-.415	-.309	-.334	-.246
	(.113)***	(.102)***	(.097)***	(.124)**	(.109)***	(.104)**
Quadratic	-.404	-.433	-.447	-.308	-.317	-.265
	(.121)***	(.110)***	(.104)***	(.128)**	(.113)***	(.108)**
Cubic	-.394	-.398	-.393	-.293	-.297	-.211
	(.120)***	(.112)***	(.106)***	(.131)**	(.118)**	(.113)*
Quartile	-.399	-.390	-.395	-.304	-.283	-.204
	(.120)***	(.112)***	(.106)***	(.130)**	(.119)**	(.113)*
Municipalities X Time	2102	2973	3679	2101	2973	3672
Municipalities	301	426	527	301	426	527
Panel B : Russian - Prussian Border						
Linear	-.031	.117	.098	-.198	.284	.251
	(.139)	(.130)	(.119)	(.153)	(.147)*	(.128)*
Quadratic	-.078	.099	.114	-.244	.264	.253
	(.146)	(.132)	(.120)	(.159)	(.149)*	(.131)*
Cubic	-.080	.090	.128	-.247	.250	.181
	(.144)	(.132)	(.124)	(.159)	(.150)*	(.134)
Quartile	-.112	.044	.134	-.273	.217	.147
	(.144)	(.134)	(.124)	(.158)*	(.152)	(.136)
Municipalities X Time	1442	2135	2898	1442	2135	2894
Municipalities	206	305	414	206	305	414
Geographic Controls	yes	yes	yes	yes	yes	yes
Socio-Economic Controls	yes	yes	yes	yes	yes	yes
Sample	rural	rural	rural	rural	rural	rural

Notes: Robust and clustered at the municipality level standard errors are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level and * at the 5%. Columns (1) to (3) - the dependent variable is the 6th grade low-stake exam score; Columns (4) to (6) the mathematics and science 9th grade high-stake exam score. Table presents estimates of the coefficient β from the regression (1) of the dependent variable on the partition dummy D, which equals 1 for the former Russian areas and 0 for either the former Austrian (Panel A) or Prussian (Panel B) territories, and the set of socio-economic variables explained in Table A1. The regressions use 50 km (columns (1) and (4)), 75km (columns (2) and (5)) and 100km (columns (3) and (6)) bandwidths.