How history matters for student performance: lessons from the Partitions of Poland

Journal of Comparative Economics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Comparative Economics

journal homepage: www.elsevier.com/locate/jce

How history matters for student performance. lessons from the Partitions of Poland☆

Paweł Bukowski

Centre for Economic Performance, The London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, the United Kingdom

ARTICLE INFO

JEL classification:
N30
I20
O15
J24

ABSTRACT

This paper examines the effect on current student performance of the 19th century Partitions of Poland among Austria, Prussia and Russia. Using a regression discontinuity design, I show that student test scores are 0.6 standard deviations higher on the Austrian side of the former Austrian-Russian border, despite the modern similarities of the three regions. However, I do not find evidence for differences across 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 education 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.

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, historical investments in public goods and property rights have been shown to affect students' current education attainment, the provision of schools and literacy levels (Banerjee and Iyer, 2005; Huillery, 2009; Iyer, 2010). However, we know less about the mechanisms underlying these long-run consequences of institutions, and if and how they depend on social context. It has been argued, for instance, that universal schooling might level the historical differences in education outcomes (Dell, 2010). In this paper, I show that two historical parts of Poland, which had similar past education systems and provision of public education, have - relative to a control region - very different long run effects on current student performance. I focus on social norms toward local schools as a channel of persistence (Akerlof and Kranton, 2010; Sakalli, 2018). I also argue that the interaction between national identity and institutions has 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 Fig. 1) as a laboratory to investigate how history matters for student performance. Comparisons of geographic characteristics and the historical literature have

☆ I thank Sascha O. Becker, Volha Charnysh, Gregory Clark, Tomas Cvrcek, John S. Earle, Irena Grosfeld, Hedvig Horvát, Gábor Kézdi, Jacek Kochanowicz, Attila Lindner, Sergey Lychagin, Christina Romer, Ruth M. Schüler, Tamás Vonyó, Jacob Weisdorf, Agnieszka Wysokińska, Noam Yuchtman, the two anonymous referees, 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 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 and from the Economic and Social Research Council grant number ES/M010341/1. All errors are mine.

E-mail address: p.bukowski@lse.ac.uk.

https://doi.org/10.1016/j.jce.2018.10.007
Received 19 September 2017; Received in revised form 11 October 2018; Accepted 30 October 2018

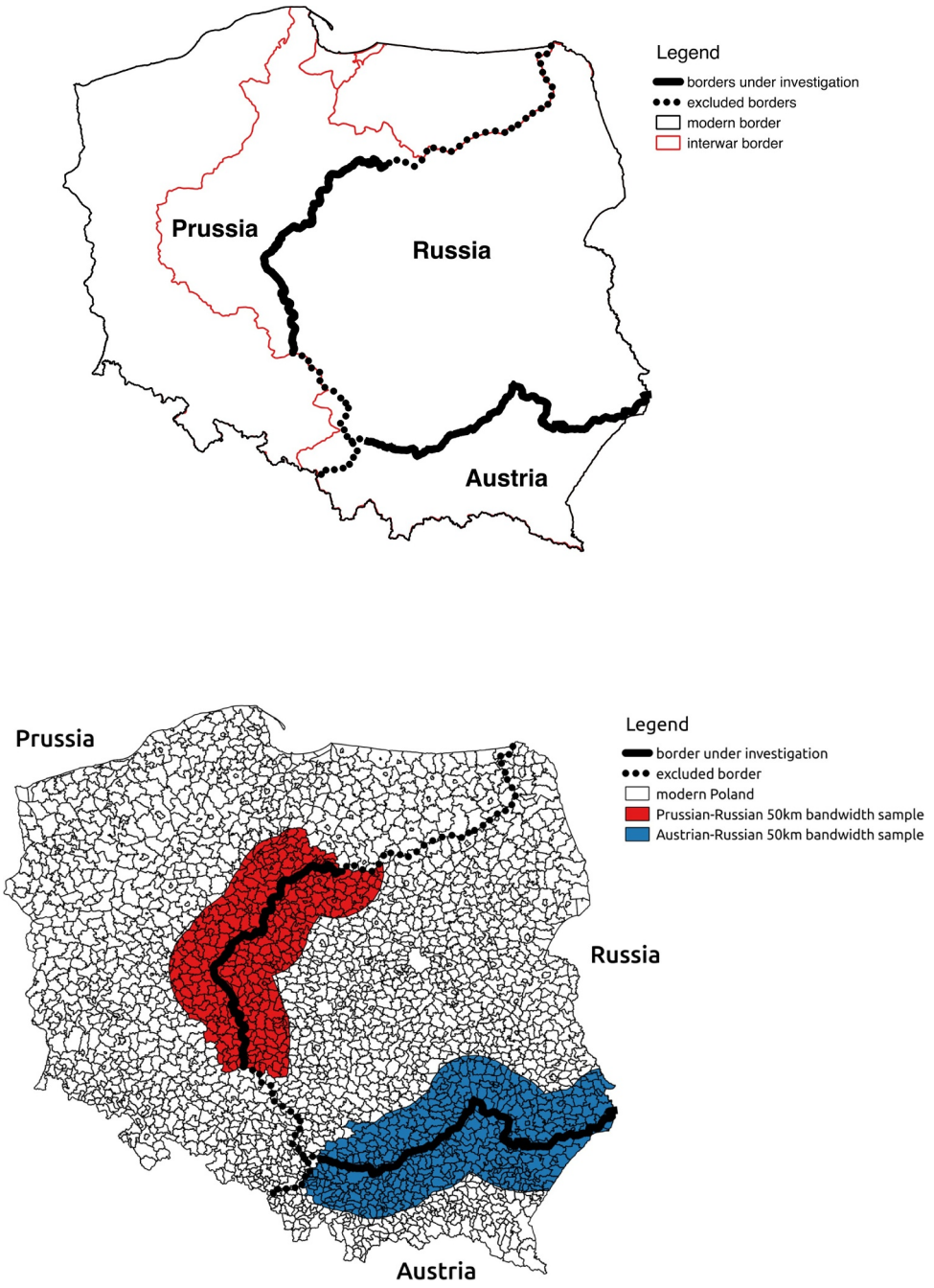

Fig. 1. The Partitions of Poland 1815–1918. Note: the maps show the borders of the 19th century Partitions of Poland layered on the modern map of Poland. The upper map also shows the borders of interwar Poland. The lower map shows the sample under investigation, based on the 50km bandwidth around the former borders. The borders under investigation are marked by the solid line, the excluded borders are marked by the dashed line. Silesia and Eastern Prussia are excluded. Source: own visualization based on GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

concluded that the former borders between the empires were not drawn to reflect any pre-existing socio-economic, historical, geographic or ethnic divisions (Wandycz, 1974; Grosfeld and Zhuravskaya, 2015; Becker et al., 2016). Consequently, I will argue that the Partitions of Poland provide an exogenous variation in institutional heritage in modern Poland. The three partitions differed significantly; However, in terms of education systems, the Austrian and Prussian institutions were similar as the former was copied from the latter (Lamberti, 1989; Cohen, 1996). The Austrian and Prussian systems were financed from local taxes, had compulsory elementary and optional secondary education, and shared similar curricula and pedagogical methods. The provision of public

education was comparable in the Austrian and Prussian partitions. The Russian education 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 education and legal systems.

Using a regression discontinuity design, I compare the test-measured performance of students in municipalities on the two sides of the former border between Austria and Russia. I show that municipality-average test scores on the Austrian side are 0.6 standard deviations higher. However, I do not find evidence for differences across the Prussian-Russian border. These results provide evidence that history matters in the long run and are consistent with other studies documenting the long-lasting effects of historical heritage (e.g. Acemoglu et al., 2001; Basten and Betz, 2013; Dell et al., 2018).

There are many potentially relevant channels through which the Partitions have affected the current student performance, but given the complexity of the treatment it is impossible to identify all of them. I focus on a social-norm channel, which has been underlined in general studies (Karaja, 2013; Grosfeld et al., 2013; Grosfeld and Zhuravskaya, 2015; Becker et al., 2016) and in the context of education outcomes (Feir, 2016; Sakalli, 2018). In particular, I argue that people living in the former Austrian Empire have inherited positive social norms toward local schools (institutions), leading to a more intensive schooling effort and higher student performance. I provide four pieces of evidence to support this hypothesis. Although each of them is tentative and parsimonious, when taken together they suggest that a social norm is an important transmission mechanism of past institutions. Firstly, I show that the effect of the Austrian Empire is larger on low-stakes exams than on high-stakes ones, which is consistent with a norm-based model of student effort (Akerlof and Kranton, 2002). Intuitively, social norms toward local schools matter more for low-stakes exams 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. Thirdly, I show that the Austrian Partition has had a positive and large effect on kindergarten attendance that cannot be explained by the historical supply of kindergartens. Finally, I investigate whether the Partitions can have persistent effects on political outcomes, which are also affected by social norms. The results show that people living in the former Russian partition are less likely to vote in central elections and those in the former Austrian partition voted more in favour of the EU accession - a proxy for the acceptance of foreign power. At the same time, I show that other channels, in particular skill-biased migrations and urbanization, are unlikely to explain my results. The historical migration patterns do not show any strong selection for in- or out-migration to or from the Austrian partition. To evaluate present-day migration, I show that a hypothetical modern extreme skill-biased migration scenario cannot explain the effect of the Austrian Partition. The Austrian partition is more urbanized and has a higher population density, however, using various methods, I show that accounting for this channel cannot explain the observed gap in student performance between the Austrian and Russian partitions.

Why social norms differ in the Austrian and Prussian partitions is puzzling given that the former was not economically superior over the latter and that the two Empires had almost identical education systems and similar provisions of public education.[2] I argue that the differential legacy of Austria and Prussia originates from the different interactions between education institutions and Polish identity. While the Prussian state used education mainly to Germanize Poles (e.g. through insisting on German as the language of instruction), the Austrian state used education to support Polish identity (e.g. through allowing Polish as the language of instruction) (Lamberti, 1989; Cohen, 1996).[3] Because of the historical attitudes of the Polish population toward the education systems, positive social norms toward education may have been more likely to emerge in the Austrian partition. These could then have been transmitted through generations and still affect student and parental effort. Consistent with this hypothesis, Steele and Aronson (1995) and Akerlof and Kranton (2010) have provided theoretical and empirical arguments 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 and identity. Using data on 19th century education outcomes in Austria and Prussia, I correlate the historical elementary school enrolment with the current student performance. The results show that the correlation is null among municipalities from the former Austrian partition, but strongly negative from the former Prussian partition. These estimates are robust to the inclusion of geographical and socio-economic covariates, yet they might not be causal. However, assuming that the remaining bias is the same in both regions, the historical expansion of the education system has a 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, given a *positive* interaction between identity and institutions in the Austrian Empire. Hence, the social norms affecting student performance have been more likely to emerge in municipalities with a larger attachment to the historical Austrian education system. Alternatively, because of a *negative* interaction between institutions and identity, more intensive historical exposure to Prussian education has led to a stronger opposing social norm toward the education system. This norm leads to a lower schooling effort and thus decreases the performance of students.

Overall, the contributions of this study are threefold. Firstly, I show that history matters for student performance and it accounts for a sizable gap in education 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.

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

[2] Consistently, Michalopoulos and Papaioannou (2013) and Grosfeld et al. (2013) have shown that not all institutions influence social norms or 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).

The studies that are closest to mine are Grosfeld and Zhuravskaya (2015), Wysokinska (2011) and Becker et al. (2016).[4] Grosfeld and Zhuravskaya (2015) found 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 or trust in government institutions. Consistent 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. Grosfeld and Zhuravskaya (2015) showed that a historical institution affects behaviour differently in different domains. In contrast, I show that it affects behaviour 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. (2016) suggest that, among the Central-Eastern European countries, the Hapsburg Empire is associated positively with trust toward local state and negatively with the acceptance of corruption. In addition, Backhaus (2018) showed that the gaps between the partitions in the *quantity* of education (elementary school enrolment and illiteracy levels) had disappeared by 1961. In contrast, I show that the gaps in the *quality* of education have persisted until today. All the studies mentioned use a regression discontinuity method.

My results are partially consistent with Herbst (2004) and Herbst and Rivkin (2012), who analyzed determinants of the distribution of exam scores in Poland. In particular, they regressed the exam scores for all municipalities in Poland on a set of modern-day control variables and the partitions dummies. They found that, relative to Warsaw, the dummy for the former Austrian part had the largest magnitude, and for the Prussian Empire the lowest. However, they did not use a regression discontinuity design and the current covariates are likely to be endogenous resulting in biased estimates (Angrist and Pischke, 2008). Moreover, the authors did not empirically identify the channels of persistence of the Partitions of Poland.

This paper is organized as follows: In Section 2, I present a historical overview of the Partitions of Poland and look in detail at the education system in each Empire. In Section 3, I describe the data and research methodology, and show the effect of the 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 the situation of the Poles in the 19th century education systems in Prussia, Austria and Russia. For readers uninterested in historical details, it is sufficient to read the summary at the end of this section. The summary also contains information about the modern education system in Poland.

The Partitions of Poland took place in three parts, during the second half of the 18th century and removed Poland from the map of Europe for 123 years, until the end of World War I.[5] The first annexation of the Polish lands by the Russian Empire, the Kingdom of Prussia and the Hapsburg Austria took place in 1772. Poland lost almost one-third of its territory and 4.5 million inhabitants. In 1793, Prussia and Russia conducted the second partition, and in 1795 all of the 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 Polish lands 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 return the situation how it had been before the Napoleonic Wars.

A 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.[6] During the first decades following the Congress, the Russian and Prussian administrations were not systematically oppressive toward Poles. The Congress Kingdom and the Grand Duchy of Poznań- newly created states controlled by Russia and Prussia respectively - experienced some level of freedom, which gave Poles hope that independence was within their reach. In the Congress Kingdom this lasted until the unsuccessful uprising against Russia in 1830, after which Poles were repressed and Russified[7] until the end of World War I. In Prussia, the situation of the Poles worsened in the 1870s, when Otto von Bismarck introduced *kulturkampf*.[8] Differently from the other partitions, the Poles under the Austrian occupation had relatively less freedom during the first part of the 19th century, but this changed after 1867 when the Austrian administration adopted a more tolerant and multicultural approach, language freedom being a significant expression of this. Polish was the official language of the Galician administration[9] 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.

[4] For the non-economic literature see: Hryniewicz (2003), Kowalski (2000), Grabowska (2004), Zarycki (2007), and Chuminski (2008).

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

[6] In this section, the "Austrian Partition" refers to the areas which correspond to the historical region of Galicia, that is, the southern parts of modern Poland and the western parts of modern Ukraine. The Prussian partition is defined as a territory of the pre-Partition Poland, which was governed by the Prussian Empire after 1815. Cities such as Wrocław or Gdańsk are not within the Prussian partition, as they were not part of Poland before 1815. The Russian partition corresponds to the central areas of modern Poland, Lithuania and parts of modern Belarus and Ukraine. In the empirical part of this paper, the definition of each partition is limited only to the territories within modern Poland (see Fig. 1).

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

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

[9] Galicia was part of Poland and Ukraine, which were both under Hapsburg rule.

In terms of the socio-political situation, the Prussian and Austrian states were more favourable to the self-organization of their Polish populations. The two Empires introduced a bureaucratic system with a strong administrative ethos (Gillis, 1971; Becker et al., 2016). The Prussian state was a state of law and, even though the administration discriminated against Poles, grass-root institutions, such as agricultural societies, credit institutions, reading rooms, newspapers and education circles, helped support the economic activity of Poles and protect their identity. Similarly, in Austria, ethnic tolerance allowed the creation of institutions spreading and preserving Polish culture. Two universities in Galicia, the Jagiellonian University in Cracow and Lviv University, played important roles in the development of Polish intellectual life. All of this was in contrast to the situation for 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 three Partitions displayed different levels of economic development. The best economic situation was in the Prussian part, where the authorities carried out many reforms, including the abolition of serfdom, agriculture, rather than industry, being the main driver of economic progress. In the Russian partition, it was industry that developed the most, with clusters of the textile industry being created in Łódź and Białystok. Warsaw became a modern city with sewers, streets, gas lighting, and a power plant switchboard. The delayed abolition of serfdom reforms during the second half of the 19th century contributed to the relative backwardness of agriculture in the Congress Kingdom. However, the worst economic situation was in the Austrian part. Before the end of the 19th century, Galicia had not been industrialized, and agriculture was underinvested and parcelled. Consequently, people experienced one of the worst poverty rates in the Hapsburg Empire, and at the beginning of the 20th century, over two million Galicians emigrated to escape severe economic conditions. However, there were no clear migration patterns with respect to education (see Section 4.2 for more details).

In 1914, the Partition empires turned against each other and placed Polish lands at a centre of the four-year conflict. World War I had tragic consequences for the Polish population and economy, but it also led to the unification of Polish lands in 1918. This new country faced a number of immediate burning challenges: the massive destructions and human losses of the Great War paralysing economic activity, chaotic and radical political scenery, social and ethnic tensions, massive unemployment and large strikes, high inflation and rural poverty amongst other. The main task was integrating the country, harmonizing the various institutions inherited from the empires including their education systems (Wolf, 2007). In addition, a series of economic downturns during the 1920s and the Great Depression hampered modernization of the country. Polish politicians and institutions became more extractive, as the system drifted from a democratic parliamentary republic towards an authoritarian presidential republic.

During World War II Poland was occupied by Germany and experienced the largest war losses. Approximately 17% of the 1939 population was killed and 62% of the national wealth was destroyed. The post-war reality, however, brought even more dramatic changes. As a result of decisions made during the Yalta conference, the country lost its eastern territories in exchange for formerly eastern parts of Germany, triggering a massive population transfer. Similarly, as after World War 1, the biggest task was to integrate the new lands and provide a smooth start for the repatriates. At the same time, a communist system with a centrally planned economy was introduced. On the one hand, the four decades of communism in Poland brought the expansion of education, along with unification, industrialization and modernization of the country. On the other hand, it led to the erosion of democratic institutions, profound economic crises and divergence from the West. Popular dissatisfaction resulted in massive protests and the emergence of the ”Solidarność” movement. In 1989, the communists were forced to organise the first (partially) free elections, which were won by the democratic opposition. After 1989, Poland introduced a market-based economy with a parliamentary democracy. The last two decades have witnessed impressive economic growth and convergence with Western Europe. These changes accelerated after Poland joined the European Union in 2004.

In the following subsections, I examine the 19th century education systems of the Polish minorities in Austria, Prussia and Russia.

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. 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 achieved the universal aim of the general intellectual development of society. 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 education was based on local taxes and municipal school boards managed school operation (Cinnirella and Schueler, 2016).

Despite its modernity and universal character, until 1870, the elementary school fell within the domain of religion (both Protestant and Catholic). Most schools were confessional, and religion was the main subject in Prussian curricula. Although the state sought to promote secular and nation-oriented[10] education, it was afraid that taking too much power from the Catholic or Protestant church would motivate them to create competitive networks of private schools. During the 1870s, Adalbert Falk - the Minister of Ecclesiastical Affairs - implemented a set of secularization reforms. These included limits on the church's influence, and the professionalization and secularization of the school inspectorate. Yet the impact of the reforms was limited, as the clergy retained a strong position.

Nevertheless, the reforms turned out to be important in the Polish context. The Catholic Church[11] helped to cultivate Polish

[10] Understood as the German nation.

[11] The Protestant church was also affected, but because of its special role in the Prussian state, to a much lesser extent than the Catholic.

national identity more than any secular movement. Consequently, the policy of secularization was carried out more consistently on 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 particular, it banned the use of Polish in administration and education, forbade the cultivation of Polish traditions, discriminated against Polish workers in the labour market, and deported Poles and Jews, who did not have Prussian citizenship.

The most important change for education was the language of instruction. In 1822, the Prussian state permitted the use of Polish in the eastern regions with significant Polish populations. 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, 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 religion classes and final exams.[12] When, in 1901 Polish, was banned completely, students and parents in Września started to protest. Soon this 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 called for a revision of the ban, the Ministry of Education retained the policy, which (Lamberti, 1989, p.109) has described as "breed[ing] germanophobia and a repugnance for the school in Polish families". But the enforced language of instruction was not the only reason why Polish parents opposed the education system.

Education inequality and feelings of unfairness were other reasons. The introduction of German as the language of instruction implied that teachers had to teach in a language in which they did not always have the required proficiency. Moreover, students from exclusively Polish speaking families had to first learn German, leaving less time for other classes. Finally, Polish schools were systematically under-financed compared to German ones (Cinnirella and Schueler, 2016). The average student-teacher ratio in the Polish lands was 93:1, while in the rest of Prussia it was 60:1 (Lamberti, 1989, p.129). All these translated into a lower quality of the Polish schools[13] and raised feelings of unfairness among Polish parents.

Moreover, the situation for teachers was ambiguous. As pointed out by Lamberti (1989), during the strikes, Polish teachers took the side of the Prussian administration because they were afraid of losing their jobs. This in turn led to acts of hostility toward these teachers: (Lamberti, 1989, p.146) describes how "the Polish press rebuked the teacher for currying favour 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". Polish parents not only distrusted and clashed with the institution of the elementary school, but also with its personnel.

Finally, as mentioned above, the education reforms aimed against the church were executed more rigorously on the Polish lands. The most profound were the introduction of interconfessional schools[14] and the secularization of the school inspectorate. From the very beginning, the Polish population viewed this innovation with distrust. As Lamberti claims, "[t]he 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, inequality, the role of teachers and interconfessional education all motivated 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 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 compulsory schooling. As such, the Prussian education system combined effective institutions alongside its set of anti-Polish regulations.

2.2. The Austrian education system

The Austrian education system was to a large extent a copy of the Prussian model (Cohen, 1996) . As far back as 1781, Joseph II had established the principle of mandatory primary education, however, until 1848 the education system mainly served as training reserved 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 reforms that modernized the education system. The strongest adherent to and executor of the reforms was the faction of German Liberals in the Austrian Parliament, who based their ideas on the Prussian model. Although delayed by a few decades, the amendments paralleled the developments in Prussia. Local tax-funded elementary schools (equivalent to Prussian schools, to which attendance was obligatory until 14) were fully introduced following the 1867 reform and the General Primary School Law of 1869. Also, secondary and higher education were modelled on the Prussian system (including the curricula).[15]

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, this was not so from the beginning. During the first part of the

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

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

[14] Interconfessional schools (also called mixed) gathered students from different religious groups.

[15] Gary B. 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).

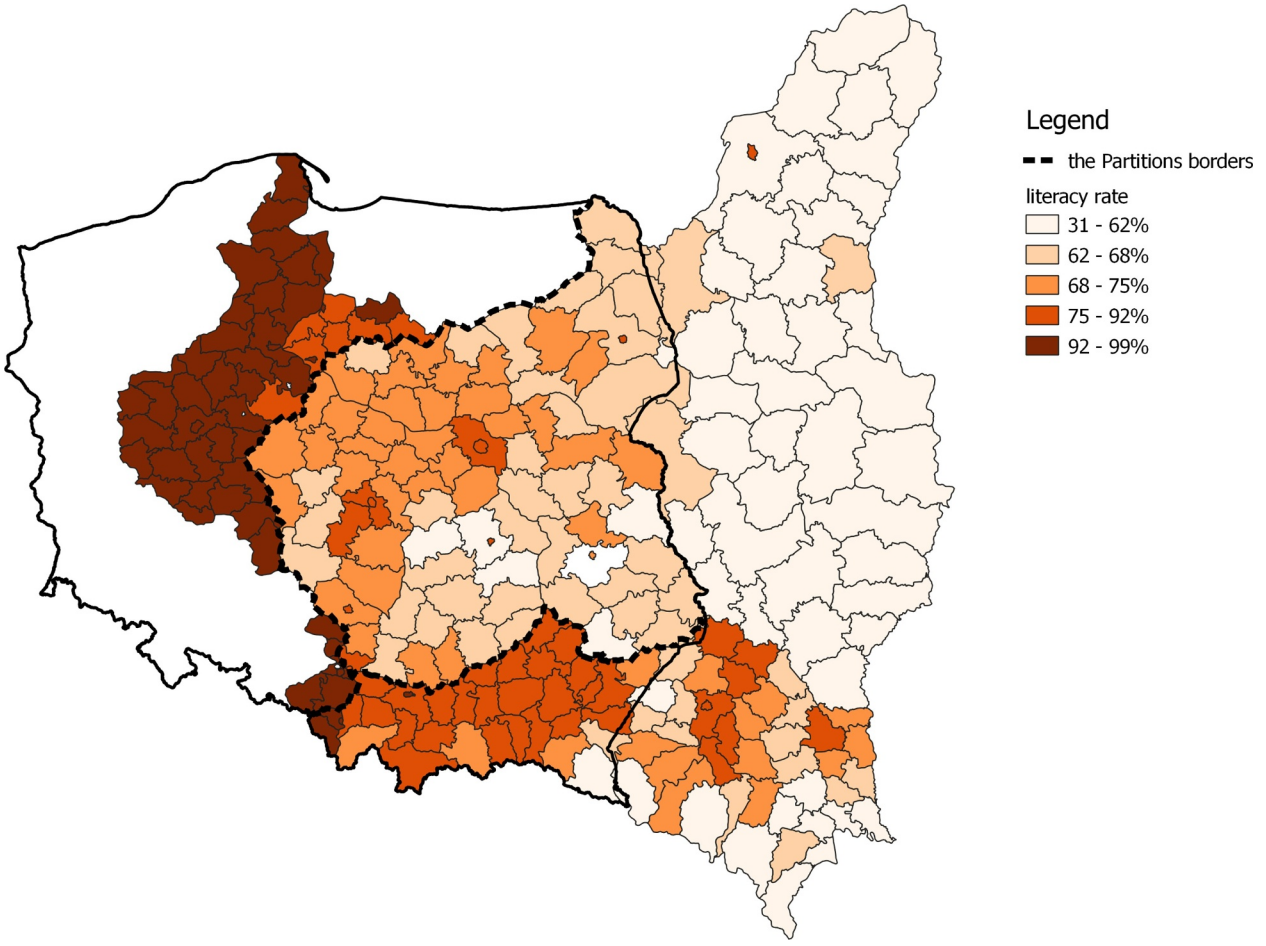

Fig. 2. Literacy levels in Interwar Poland (1931). Note: the map shows the county-level literacy levels in 1931. Literacy means "able to read and write". The borders of the Partitions of Poland are marked by the dashed line. The modern borders of Poland are marked by the solid line. Source: author's computation using the Census of 1931. Own visualization based on GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

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 local elites favoured public education only if it was in German. This changed after 1867, when the second wave of reforms extended the Polish language of instruction to secondary and higher education.

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 from previously uneducated groups resulted in numerous grass-root education initiatives. The non-German speaking ethnic groups and Jews had greater aspirations toward education than Germans. Also new lower-middle classes, for instance, the 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 agriculture was backward and extensively parcelled. The beginning of the 20th century saw rapid growth in elementary and secondary education in Galicia. The share of elementary school students in the population almost tripled between 1880 and 1910 (GUS, 2003) (see Table 2) and the secondary enrolment ratio increased by 120% (in the German speaking lands it increased by 52%) (Cohen, 1996).[16] As pointed out by Cohen (1996, p.257), "by 1910 the Polish speaking share of Austrian enrolments significantly exceeded the Polish speaking share of the Austrian population". There was also much popular and political pressure to open advanced education to children from poorer strata. At the same time, however, literacy levels and schools' attainment were still lower there than in the Prussian Partition or other parts of the Austrian Empire.[17]

[16] The secondary enrolment analysed per thousand inhabitants in the Polish speaking lands of the Hapsburg Empire was: 1880 - 2.74, 1890 - 2.78, 1900 - 3.77, 1910 - 6.05; in the German speaking lands it was: 1880 - 3.88, 1890 - 4.04, 1900 - 4.61, 1910 - 5.88 (Cohen, 1996, p.141).

[17] Cohen claims: "In the 1870s and 1880s the majority 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 the total population in Austria older than 11 was literate while in Galicia this number was 58%"(1996, p.64).

The Poles living in Galicia saw education as the principle means for preserving national identity and improving their material conditions. Even though law enforcement and the quality of institutions were not as good as in the Prussian Empire, the education system created positive relationships between schools and citizens.

A remaining question, the answer to which is beyond the scope of this paper, is why the Prussian and Austrian Empires implemented different ethnic policies from the second half of the 19th century. The Austrian state was weaker and the territory more ethnically fragmented and, consequently, a policy of Germanization could lead to social unrest. In fact, it was the revolutions of 1848 that contributed to the reorientation of the ethnic policy of the Hapsburg Empire. Conversely, the Prussian Empire was relatively homogeneous, the state stronger and thus the Germanization policy was more likely to be successful.

2.3. The Russian education system

Peter I and Catherine II created an education system whose core remained largely unchanged until the end of the 19th century. Besides much investment in universities and growing numbers of enrolled elementary school students, members of the ruling class did not accept the Humboldtian approach to education. Sergei Uvarov, the Minister of Education (1831–1849) during the rule of Nikolai I, may be the best example of one such dignitary. He laid the foundations for the modern and high-quality higher education in Russia[18], 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 that ”the lower classes had to be protected from too much knowledge” (Kassof, 2004).

Another problem was the chaotic organization of the school system. There was no obligatory schooling and the Ministry of Education did not control the network of schools.[19] Lacking central organization, the system was characterized by class-based duality, with separate curricula for students from the upper and lower strata. Consequently, the illiteracy levels were high: in 1917, only 70% of the urban and 30% of the rural population could read and write.[20]

The situation was especially bad in the Polish lands (the Congress Kingdom). The lack of education institutions was accompanied by intensive Russification and repression of Poles[21] (Chubarov, 2000). 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 Tsarist Russia, especially at the end of the 19th century. This becomes undoubtedly true once we think about the general situation of 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. Compared with other parts of Poland, its development of education could have been wasted.

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 Prussian authorities allowed the self-organization of Poles, which contributed to the preservation of Polish culture, which otherwise may have been threatened by Germanization. In the Russian partition, industrialization led to modernization and the development of cities. But weak law enforcement and the anti-Polish orientation of the Tsarist policy undermined the position of 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, from the 1860s, significant autonomy was given to the Polish population.

In terms of education systems, the Austrian and Prussian institutions were very similar as the former was copied from the latter. They were both 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 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, Poles under Prussian rule opposed the education system and were hostile toward school personnel (especially teachers). Remarkably, massive school strikes were organized by Polish parents, the largest one taking place in 1901 when 70 thousand Polish students refused to go to school. The Russian education system, in turn, practically did not exist in the 19th century.

The differences in the education outcomes between the three partitions are documented in Table 2. School enrolment in the Prussian part in 1864 was as high as 93%. In the Austrian part enrolment was significantly lower throughout the 19th century, but it had converged with 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 1910s was practically the same in the Austrian and Prussian partitions: on average, in the former, there was one school per 13 km^2, and in the latter, one school per 10 km^2. However, in the Russian partition, there was one school per 27km^2. After Poland gained independence in 1918, the literate population was as low

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

[19] Kassof (2004) estimates that ”sixty-seven different types of primary schools [existed] in Russia in1914″.

[20] 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”.

[21] 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.

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 Poles	Low	High	High
Quality of Bureaucracy	Low	High	High
Education systems			
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 education outcomes.

Partition/Year :	1840′	1850′	1860′	1870′	1880′	1890′	1900′	1910′
Elementary school enrolment								
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 who can read, aged 9 or less								
Russian	-	-	18%	-	-	41%	-	-
Austrian	-	-	-	-	-	-	-	69%[a]
Prussian	-	-	-	-	-	-	-	95%[b]

Notes: a: share of population aged 11 or less. 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).

as 35 percent in certain formerly Russian lands, whereas in the former Prussia, it was above 90 percent. The literacy levels in 1931 are depicted in Fig. 2. Regions in the West had the highest levels of literacy; these were moderately lower in the South (except for the presently Ukrainian parts), and the lowest levels were in the Central and Eastern parts of Poland. These differences were smoothed out after World War II, when 8-year education became obligatory in Poland (Meissner and Majorek, 2000; Backhaus, 2018). Yet, social norms toward education could not be so easily smoothed.

Between 1999 and 2017, the Polish comprehensive and compulsory education system consisted of 6 years of elementary school, followed by 3 years of lower-secondary school (*gimnazjum*). Admission to comprehensive schools was based on catchment areas, which means that each student living within an area has a right to attend a local public school. However, parents may request an alternative school, but its principal has the right to reject applications. After the 9th grade, students enter the tracking and non-obligatory part of education. They can choose a track (academic, mixed or vocational) and apply to any high schools, but admission is not guaranteed.

Table 3
Descriptive statistics.

Border:	Russian-Prussian					Russian-Austrian				
Variable/Partition:	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	0.146	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	0.926	0.47**	4.89	1.88	4.16	1.31	0.73***
Kindergarten attendance (2011 in %)	58.6	14.7	58.2	13.4	0.45	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	0.56
Class size (2009)	16.2	2.6	14.6	2.7	1.6***	15.3	2.4	13.8	2.62	1.51***
Elementary schools per km^2 (2011)	0.036	0.015	0.039	0.016	0	0.084	0.038	0.042	0.02	0.05***
Secondary schools per km^2 (2011)	0.014	0.007	0.014	0.007	−.003	0.038	0.038	0.017	0.01	0.021***
Expenditure per capita (2011 in PLN)	3172	789	3111	617	60.5	3204	735	3079	685	125
Edu. expenditure 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	0.85 ***	21.2	2.02	18.6	1.56	2.63***
Unemployment(2011 in %)	9.11	0.336	10.2	0.376	−1.09*	9.28	0.311	9.39	0.297	−.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	0.26***
Number of municipalities	80		126			164		137		

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

3. The partitions of Poland and student performance

3.1. Data

My analysis draws on comprehensive municipality-level registry data on obligatory, standardized and externally graded examination scores for the period 2005 - 2011, published by the Central Examination Board of Poland. The available exam scores are from a low-stakes general 6th grade exam (taken at the end of elementary school) and a high-stakes mathematics and science 9th grade exam (taken at the end of lower-secondary school). While the former serves mainly a diagnostic purpose, the latter matters in the high school admission process and thus motivates students (and their parents) to obtain their best score. The tests are standardized at the country level and are corrected double-blindly outside school by randomly chosen professional test checkers.

Socio-economic control variables at the municipality level come from the Central Statistical Office of Poland and the System of Education Information. Geographical and climate data come from the *WorldClim.org* project (Hijmans 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. Since the descriptives are for the present period, they might reflect the effect of the Partitions of Poland. The border areas of the former Russian partition seem to have the worst socio-economic situations, as the rates of unemployment are highest, expenditures are lowest and the migration balances are negative. Importantly, the former Russian areas on the Austrian-Russian and Prussian-Russian borders are today similar (Columns 3 and 8). The municipalities that were under Prussian rule are characterized by a high share of employment in agriculture[22], a high share of people aged 0–18, positive migration balance and low levels of unemployment. The situation in the former Austrian zone is similar to the former Prussian, except for the lower importance of agriculture and a higher population number and density.

The rural borderlands of the former Austrian partition have the best education outcomes (except the number of additional classes and the level of scholarization), even though these lands do not necessarily perform better in other socio-economic indicators. Importantly, they also have higher education 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 levels of achievement, but the former have larger classes and a higher number of additional lessons. The high performance of students from the former Austrian partition is also confirmed by Figs. 3 and 4, which show the spatial distribution of the 6th and 9th grade exams in 2011, for the whole country. It can

[22] The agriculture practice on the former Prussian lands is based on large, business-oriented farms, which are not common in the rest of the country.

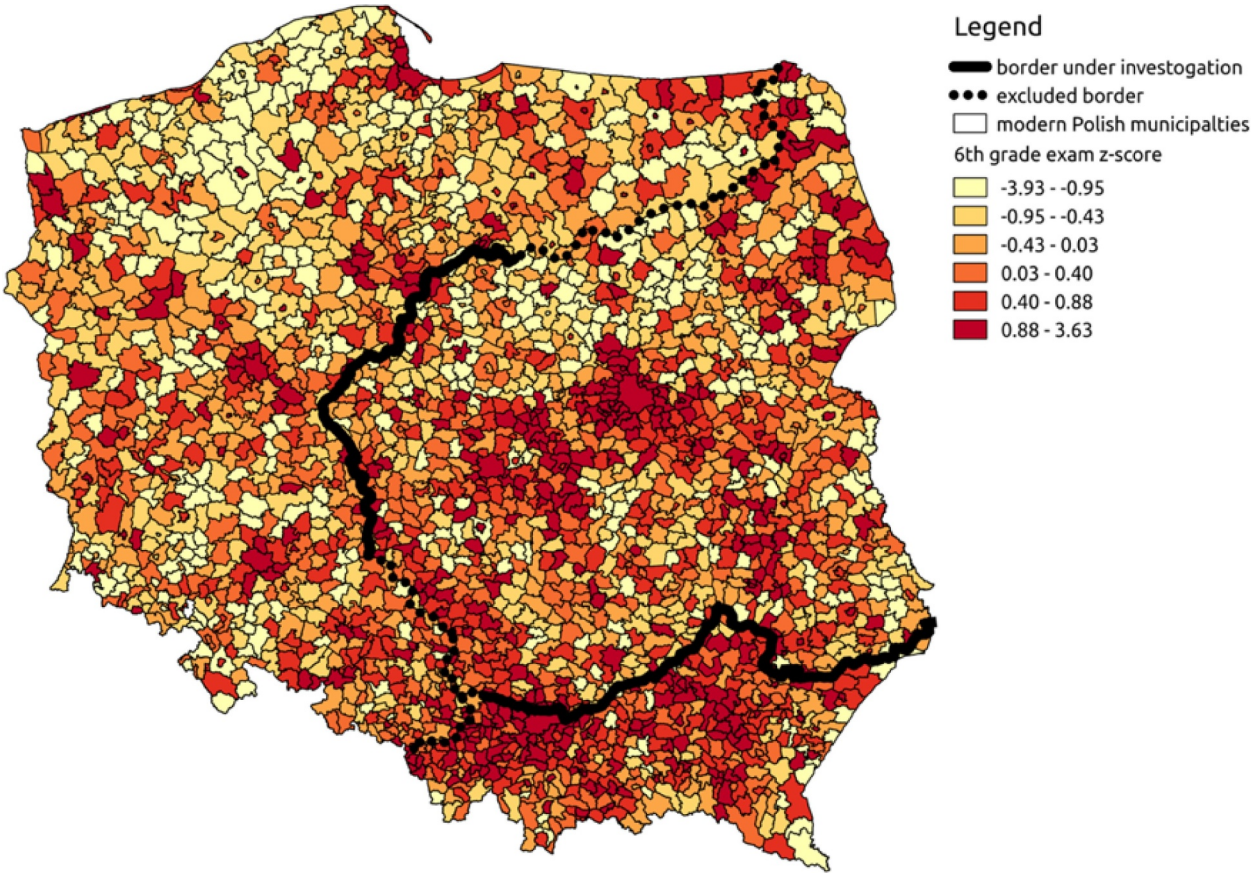

Fig. 3. The distribution of the 6th grade exam score in 2011. Note: the map shows the distribution of the 6th grade low-stakes exam score in 2011 at the municipality level. The borders of the Partitions of Poland are layered on the modern map of Poland. The border under investigation are marked by the solid line, the excluded border are marked by the dashed line. Source: own visualization based on the Central Board of Examination data and GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

be clearly seen that the territory of the former Austrian Empire houses a cluster of high-performing municipalities.

3.2. Empirical strategy

A straightforward comparison of schools in the former Prussian, Austrian and Russian partitions neglects unobserved differences between these areas, leading to biased estimates of the effect of the Partitions of Poland. To solve this problem, I follow Dell (2010) and employ a geographical two-dimensional regression discontinuity design, which evaluates the effect of the Partitions by focusing on a discontinuous jump at the borders.[23] To control for the potential confounding effects of geographical location, I narrow the analysis only to areas located close to the partitions' borders and include into the regression a polynomial of latitude and longitude. The model can be written as:

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

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 outcome variables are the standardized (Z-score) 6th grade exam score and the standardized mathematics and science 9th grade exam scores,[24] which are available from 2005 to 2011. The sample consists of municipalities that are located within a given distance to the borders (the bandwidth). I pool the data and estimate the model using a random effect estimator, as it produces more efficient estimates in the presence of individual effects. Nevertheless, in order to see whether the results are not driven by any particular year, I also reported OLS estimates for each year separately.

[23] For more about the geographical regression discontinuity design see Keele and Titiunik (2014), for general discussion about the regression discontinuity framework see Imbens and Lemieux (2008), Lee and Lemieux (2010), or Angrist and Pischke (2008).

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

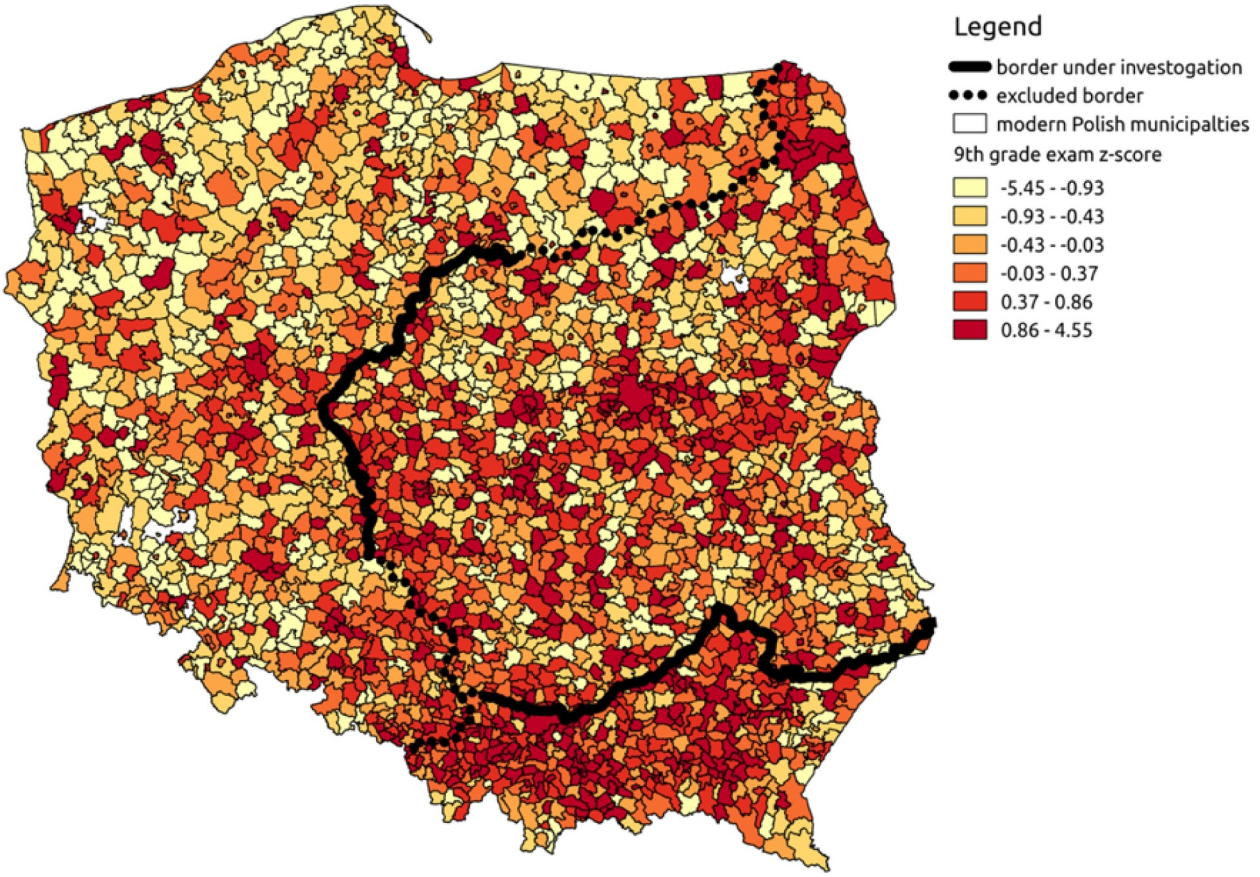

Fig. 4. The distribution of the 9th grade exam score (math and science) in 2011. Note: the map shows the distribution of the 9th grade high-stakes exam score in math and science in 2011 at the municipality level. The borders of the Partitions of Poland are layered on the modern map of Poland. The border under investigation are marked by the solid line, the excluded borders are marked by the dashed line. Source: own visualization based on the Central Board of Examination data and GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

The regression discontinuity framework requires a proper specification of the polynomial $f(location_i)$ and the bandwidth. Table A4 reports the Akaike Information Criteria, which can be used for the model selection (Lee and Lemieux, 2010). Quadratic and quartile polynomials are favoured, however Gelman and Imbens (2018) argue that high-order polynomials might be misleading in the regression discontinuity design. In the baseline regression, I use a quadratic polynomial, but I also report results for linear, cubic and quartile polynomials and, in Section 3.5, I show that the results are not sensitive to the polynomial selection. The bandwidth selection is based on the trade-off between the sample size and internal validity. For my baseline specification, I chose a 50km bandwidth, but I also report results for municipalities located at most 75km and 100km from the borders.

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 education performance are smooth at the border. There is a consensus among historians that the borders of interest[25] were not drawn to reflect pre-existing socio-economic, historical, geographic or ethnic divisions (Wandycz, 1974, p.11). Becker et al. (2016) showed that there were no significant differences between these regions in terms of their pre-Partition characteristics. Backhaus (2018) using data from the 1810 census of the Duchy of Warsaw provided evidence that there was no difference in the size of population at the Prussian-Russian border, which was established in 1815.[26] In addition, I have excluded 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.[27] Thus, I ensure

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

[26] The 1810 census of the Duchy of Warsaw did not cover the areas of the Austrian partition and thus it is not possible to compare the borderlands of the Austrian-Russian border.

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

Table 4
Geographic differences.

Dep. Variable:	Altitude (m)		Precipitation (mm)		Temperature (C°)	
	(1)	(2)	(3)	(4)	(5)	(6)
The 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)***
Mean of outcome	252	249	641	639	7.71	7.73
R^2	.58	0.58	0.93	0.93	.30	0.31
Municipalities	301	373	301	373	301	373
Panel B: Quadratic polynomial in distance						
Partition effect	8.9	6.8	8.9	6.2	0.03	0.05
(Russia = 1)	(20.1)	(18.7)	(17.4)	(16.3)	(.12)	(.11)
Mean of outcome	252	249	641	639	7.71	7.73
R^2	.30	0.30	0.37	0.36	.34	0.35
Municipalities	301	373	301	373	301	373
The 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)***
Mean of outcome	115	113	540	539	8.01	8.01
R^2	.83	0.81	0.85	0.84	.95	0.94
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)
Mean of outcome	115	113	540	539	8.01	8.01
R^2	.09	0.07	0.02	0.03	.05	0.03
Municipalities	206	302	206	302	206	302
Sample	Rural	All	Rural	All	Rural	All

Notes: Robust and clustered standard errors (at municipality level) 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 millimetres; 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.

that the observed differences between the areas of interest are due to the Partitions of Poland, not some later historical event. Overall, my sample consists of areas which experienced a similar history *before* and *after* the Partitions, were ethnically homogeneous and are now within the territory of Poland. Fig. 1 depicts the partitions' borders layered on the modern boundaries of Poland, with the solid line representing the borders under investigation, and the dashed line marking the excluded parts. By excluding Silesia, I cannot directly compare the borderlands between the Austrian and Prussian partitions. The bottom part of Fig. 1 presents the 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 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, the results based on the sample with include urban municipalities are also reported in the robustness section.

The partitions' borders under investigation were mostly set along rivers. The border between the Prussian and Russian Empires was drawn along the Drw'ca and Prosna rivers (which are small waterways), whereas half of the Austrian-Russian border was drawn along the Vistula river. Grosfeld and Zhuravskaya (2015), using a one-dimensional non-parametric regression discontinuity model, report only a small difference in altitude on the Austrian-Russian border. Yet, according to my own estimations in Table 4, Panel A shows that there are significant differences in temperature and precipitation on the Austrian-Russian border when a two-dimensional specification is used. The "jump" in altitude on the Russian side of the Austrian-Russian border is around 80 m, precipitation is higher by 30mm and temperature drops by around 0.45 C°. The magnitudes are not large, and they arise 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.[28] I control for the geographic and climate characteristics in the baseline regressions (the estimates are generally insensitive to their inclusion). In addition, at the end of Section 3.5, I show that placebo experiments run on other rivers in Poland do not show any effect on the performance of students. Overall, I find it very unlikely that these natural differences could explain education differences between the borderlands or induce other dissimilarities in culture or institutions.

[28] Consistent with Grosfeld and Zhuravskaya (2015), the one-dimensional specification in Table 4 Panel B and Panel D does not produce significant differences on either of the borders.

Fig. 5. One-dimensional regression discontinuity: the Austrian-Russian border. Note: the left panel of the graph shows the municipality averages of the standardized 6th grade exam score from 2005 to 2011 plotted against distance to the Austrian-Russian border. The right panel shows analogous graph for the municipality averages of the standardized 9th grade exam score 2005–2011. Negative distance is for the Austrian side. The graph plots the rural municipalities only. Source: own visualization based on the Central Board of Examination data.

When the border exogeneity assumption is satisfied, estimation of the discontinuous change in the outcome variable at the borders yields a causal effect of the Partitions of Poland. The channel of influence might be through social norms, migration or other processes induced by the Partitions.

3.4. Results

Figs. 5and 6 present the relationships 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 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 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-stakes exam score as a dependent variable, while Columns 5 to 8 with the mathematics and science 9th grade high-stakes exam score. The results for the rural sample are reported in Columns 1, 2, 5 and 6. Additionally, in Columns 2, 4, 6 and 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 deviations (σ)[29] and on the 9th grade exam by 0.42σ (Columns 1 and 5). All the coefficients are strongly significant. The magnitudes and economic importance of the results are comparable to the Black vs. White achievement gap in the US (Lee et al., 2007). The estimates drop to 0.54σ and 0.4σ respectively, once I add the set of geographic control variables (Columns 2 and 6). The smaller effects on the 9th grade high-stakes exam are consistent with the social norm hypothesis, which predicts that the gap between regions with different social norms will 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 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. Unlike for the Austrian-Russian border, these results show that the 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

[29] To obtain the effect of the Austrian or the Prussian Empires one simply changes the sign of the coefficients reported in Table 5.

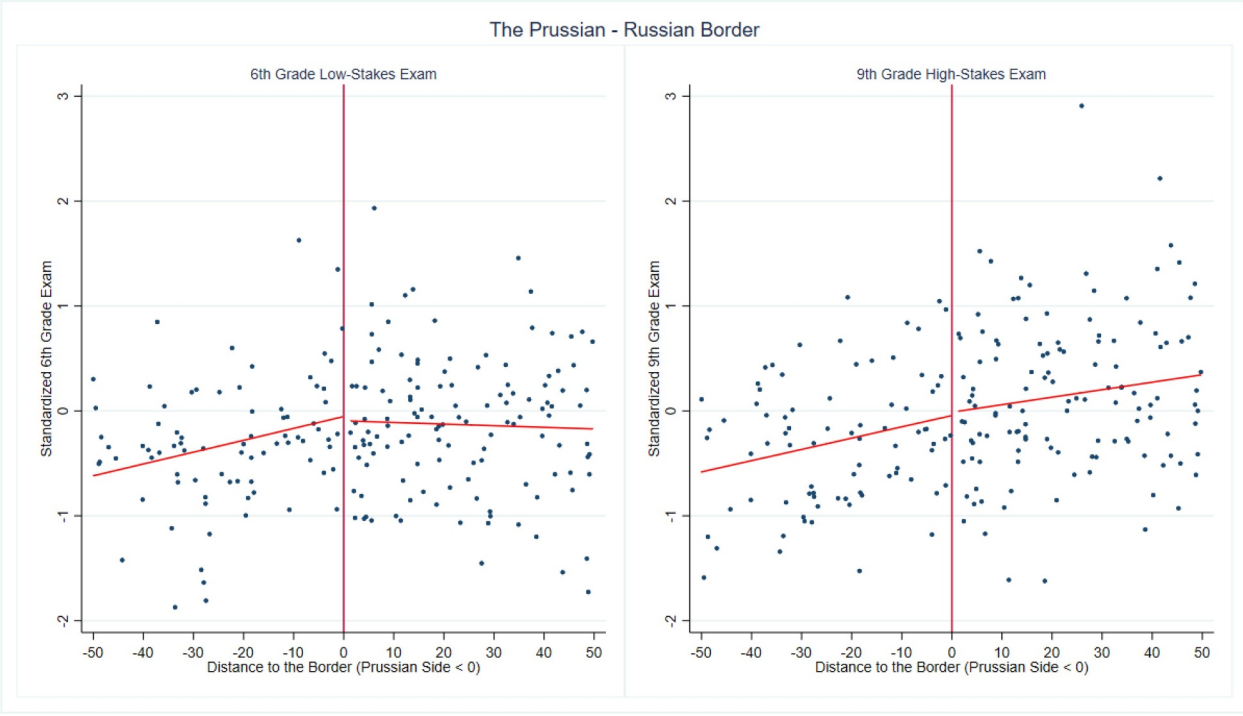

Fig. 6. One-dimensional regression discontinuity: the Prussian-Russian border. Note: the left panel of the graph shows the municipality averages of the standardized 6th grade exam score from 2005 to 2011 plotted against distance to the Prussian-Russian border. The right panel shows analogous graph for the municipality averages of the standardized 9th grade exam score 2005–2011. Negative distance is for the Prussian side. The graph plots the rural municipalities only. Source: own visualization based on the Central Board of Examination data.

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 Fig. 7, which shows 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 thus limiting the sample only to the rural areas would introduce a sample selection bias.[30] Therefore, as a first robustness check, I estimate Eq. 1 on the total sample. I also include categorical and continuous variables indicating the population size and density of a municipality. This is a less preferable sample, since there are few large cities that 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 relationship between location and 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 a one dimensional model, where $f(location_i)$ from Eq. (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.[31] I centre 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 the exception of Column 1).

The baseline results might be sensitive to the specification choices. Table A5 reports estimates of the Partitions' effects for different polynomials in latitude and longitude, along with different bandwidth choices. All regressions include the geographic

[30] Suppose that the Austrian Empire positively affected the urban population growth. Limiting the sample only to municipalities smaller than 50 thousand 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.

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

Table 5
Baseline regressions.

Dep. Variable:	6th grade low-stakes exam				9th grade high-stakes exam			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
The Russian - Austrian border								
Panel A : Quadratic polynomial in latitude and longitude								
Partitions effect	−.637	−.561	−.615	−.554	−.432	−.409	−.404	−.402
(Russia = 1)	(.119)***	(.125)***	(.107)***	(.116)***	(.116)***	(.131)**	(.104)***	(.122)***
Mean of outcome	.1	0.1	0.19	0.19	0.02	0.02	0.1	0.1
R^2	.26	0.3	0.28	0.32	0.19	0.2	0.21	0.22
Municipalities	301	301	373	373	301	301	373	373
Mun. × Time	2107	2107	2606	2606	2106	2106	2605	2600
Panel B: Quadratic polynomial in distance								
Partitions effect	−.45	−.483	−.493	−.502	−.231	−.273	−.343	−.357
(Russia = 1)	(.256)+	(.235)*	(.234)*	(.215)*	(.229)	(.217)	(.231)	(.217)
R^2	.14	0.19	0.2	0.24	0.09	0.16	0.15	0.2
Mean of outcome	.1	0.1	0.19	0.19	0.02	0.02	0.1	0.1
Municipalities	301	301	373	373	301	301	373	373
Mun. × Time	2107	2107	2606	2606	2106	2106	2605	2605
The Russian - Prussian border								
Panel C : Quadratic Polynomial in latitude and longitude								
Partitions effect	−.035	−.059	−.134	−.153	−.072	−.132	−.042	−.09
(Russia = 1)	(.156)	(.156)	(.125)	(.128)	(.172)	(.169)	(.131)	(.132)
Mean of outcome	−.22	−.22	−.15	−.15	−.04	−.04	−.03	−.03
R^2	.06	0.07	0.11	0.11	0.11	0.13	0.11	0.14
Municipalities	206	206	302	302	206	206	302	302
Mun. × Time	1442	1442	2114	2114	1442	1442	2114	2114
Panel D : Quadratic polynomial in distance								
Partitions effect	−.405	−.384	−.334	−.334	−.144	−.118	−.135	−.126
(Russia = 1)	(.237)+	(.24)	(.204)	(.203)+	(.24)	(.249)	(.188)	(.195)
R^2	.04	0.08	0.08	0.12	0.07	0.09	0.07	0.1
Mean of outcome	−.22	−.22	−.15	−.15	−.04	−.04	−.03	−.03
Municipalities	206	206	302	302	206	206	302	302
Mun. × 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 standard errors (at municipality level) 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-stakes exam score; Columns 4 to 6 - the mathematics and science 9th grade high-stakes 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, B) or Prussian (Panel C, D) territories. In addition, the regressions include a quadratic polynomial in latitude and longitude (Panel A, C) or a quadratic polynomial in distance (Panel B, D), and a set of geographic covariates (columns 2,4, 6 and 8). All the regressions use 50 km bandwidth.

controls. 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. The estimates on the Prussian-Russian border vary considerably across specifications. Importantly, the sign changes once the bandwidth is increased to 75km and 100km, which indicates that students living in the Russian zone, in fact, perform better than those from the Prussian zone. These contradict the findings from Table 5. Nevertheless, in most cases the coefficients are not significant.

The same set of specification choices is examined with the total sample (Table A6) and with the one-dimensional regression discontinuity design for the rural sample (Table A7). Both tables consistently show highly significant (except cubic and quartile specifications in Table A7) and positive effects of the Austrian Empire. The results for the Prussian Empire are similar as previously.

In order to check whether the results are not driven by any particular year, I estimate the baseline Eq. (1) by OLS on the rural sample for each year. Fig. 8 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-stakes exam and similar, but smaller, in the case of the 9th grade high-stakes exam. By contrast, the effect of the Prussian Empire is null across years and types of exam.

Next, I check whether the results are sensitive to an inclusion of a set of time-variant modern general education controls (local government expenditure per capita on education, and kindergarten and secondary school attendance) and a set of modern socio-economic controls (local government total expenditure per capita, unemployment ratio, population level, population density, school density and migration balance). Table A1 provides definitions of these variables. These covariates are endogenous, that is, they are

Table 6
Discontinuities with log of covariates as dependent variables.

Dep. Variable/Border:	Prussian-Russian (1)	Austrian-Russian (2)
Panel A : Time-variant variables		
Expenditure	.029	.092
	(.051)	(.051)
Education expenditure	0.079	.03
	(.051)	(.049)
Unemployment rate	.104	−.223
	(.082)	(.063)***
Sec. school scholarization	.056	.036
	(.040)	(.020)+
Population	−.204	−.407
	(.128)	(.124)***
Population 0–18	.011	−.054
	(.015)	(.015)***
Population density	.144	−.249
	(.091)	(.084)**
Elementary schools per km^2	.235	−.104
	(.094)*	(.084)
Secondary schools per km^2	.219	−.136
	(.104)*	(.082)
Municipalities	206	301
Municipalities × Time	1442	2105
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)
Education Value Added	0.074	0.004
	(.419)	(.298)
Municipalities	206	298
Geographic controls	Yes	Yes
Sample	Rural	Rural

Notes: Robust and clustered standard errors (at municipality level) 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 1 of *logarithms* of various dependent variables (except education 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 Russian 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.

also affected by the Partitions of Poland (see Section 4.3 and Table 6). The augmented regression, therefore, "switches off" some potential channels of influence of the Partitions on outcomes and leads to bad-control bias (Angrist and Pischke, 2008).[32] Indeed, as reported in Table A8, the effect of the Austrian Partition drops slightly, but remains significant and large. 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 exam. On the other hand, the estimates of the effect of the Prussian Partition are not significant and sometimes have an opposite sign.

The current border between voivodeships (NUTS2 administration level) overlaps almost completely with the former Russian-Austrian border. If voivodeships 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 highly decentralized (Herbst et al., 2009). A local municipality's government manages the school network of public elementary and lower secondary schools, and the role of the central government is limited to financing education and enacting general resolutions. Moreover, voivodeship administration is irrelevant for the education governance. Consistently, there is no significant effect of the Austrian Empire on municipalitys' education spending per capita (Table 6). Secondly, since the former Russian-Prussian border does not overlap with the administrative borders, I can include the voivodeship dummies in Eq. (1). None of these dummies are significant,

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

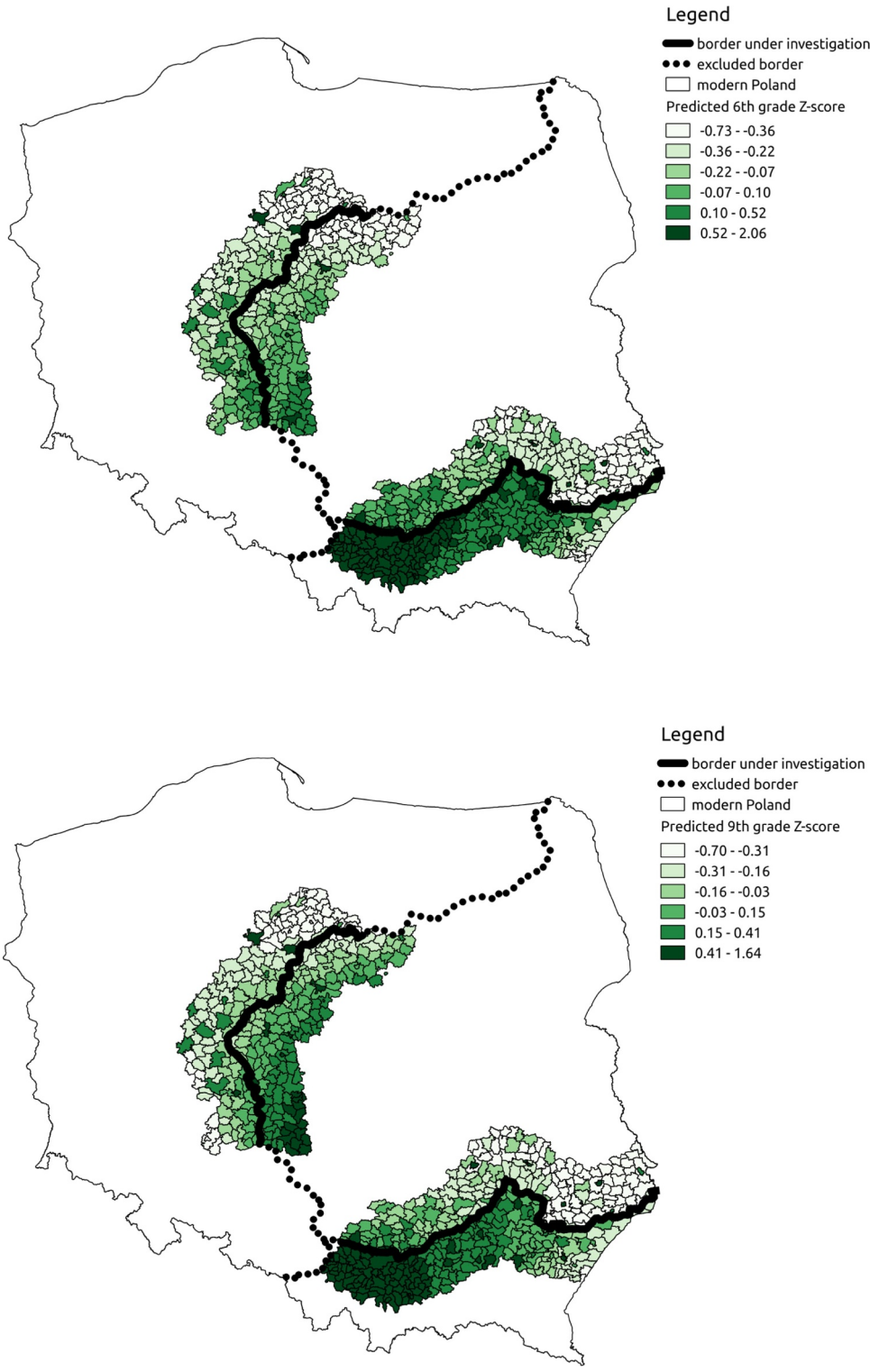

Fig. 7. Predicted levels of the 6th and 9th grade exam scores for 2011. Note: the map shows the predicted values of the 6th and 9th grade exam Z-score for 2011, based on the regressions specified im Eq. 1, using the whole sample, and including a quadratic polynomial of longitude and latitude, the partition dummy, geographic covariates and a set of population size dummies. The borders of the Partitions of Poland are layered on the modern map of Poland. The border under investigation are marked by the solid line, the excluded borders are marked by the dashed line. Source: own visualization based on the Central Board of Examination data and GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

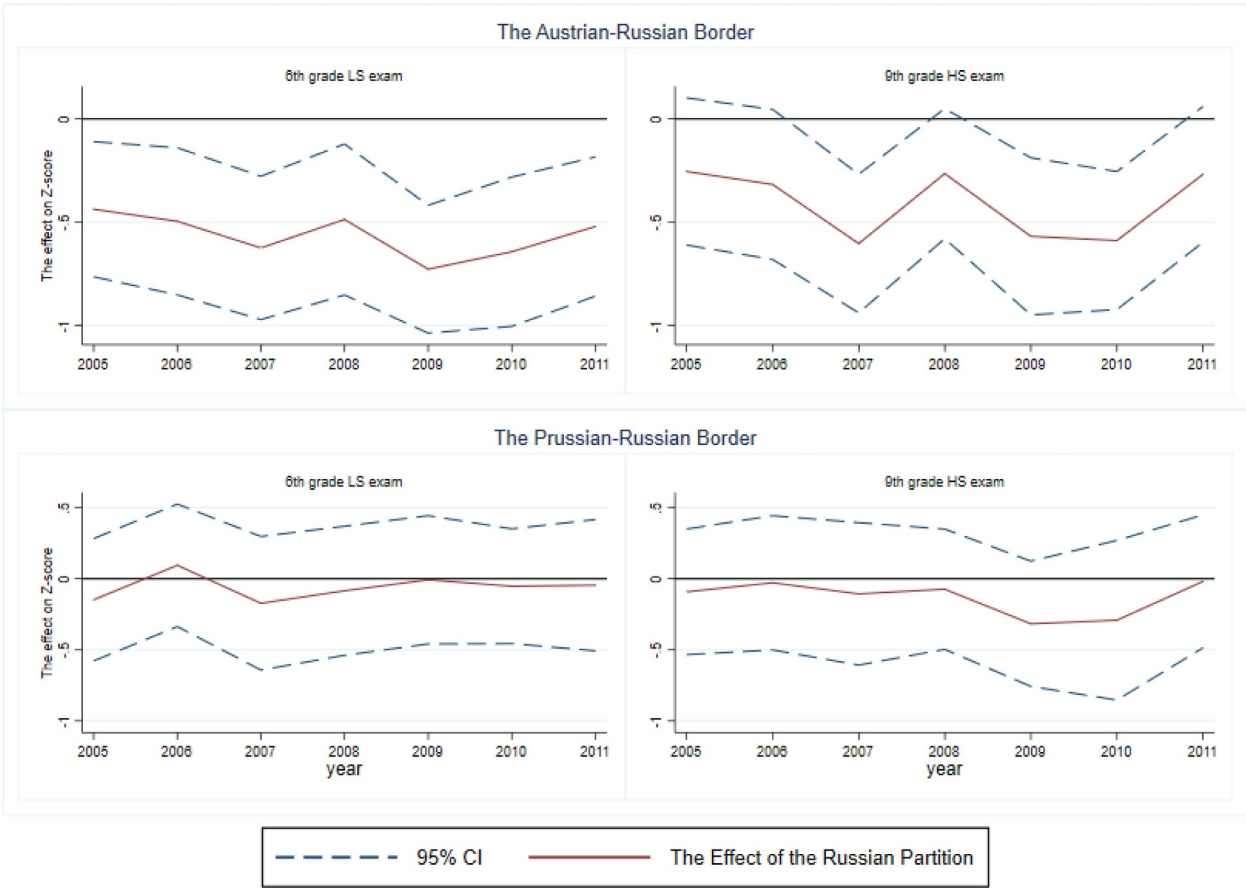

Fig. 8. The Partitions effect year-by-year. Note: the figures present the OLS estimations of Eq. 1 based on the rural sample, for each year separately, including a quadratic polynomial of longitude and latitude and geographic control variables. Only the coefficients and confidence intervals of the Partitions dummy (the variable of interest) are presented. The dummy takes value one for the former Russian areas and zero for either the former Austrian or Prussian areas. Source: author's computation using the Central Board of Examination data.

suggesting that a voivodeship 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 I estimate the baseline regression at an artificial border, which I create by moving step-by-step the actual Austrian-Russian border by 5km to the North or to the South (at most around 100 km). I define the "Russian" dummy as an area north of the artificial borders. Fig. 9, Panel A presents Z-tests[33] 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 of the artificial borders. Fig. 9, Panel B shows Z-tests. This time the actual border is not different from the other artificial borders.

As discussed in Section 3.3, the Partitions' borders were set along rivers (the Drw'ca, the Prosna and the Vistula) and there are minor differences in geographic characteristics across the Austrian-Russian border. The rivers might separate regions with, for instance, different urbanization patterns, leading to differences in student performance. In order to investigate this possibility, I run four placebo experiments, with artificial borders set on other rivers in Poland. I chose two artificial river borders with East-West orientation similar to the Austrian-Russian border. The first one goes along the Bug, Vistula and Bzura rivers; the second along the Wieprz, Vistula and Pilica rivers. Similarly, I selected two artificial borders with North-South orientation similar to the Prussian-Russian border. The first is located on the Narew, Bug, Vistula and Pilica rivers. The second along the Odra river. In all cases, the effect of the placebo "Russian" dummy was insignificant, which strongly suggests that the rivers are unlikely to drive my results (results available upon request).

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, small and changes sign across specifications.

[33] 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.

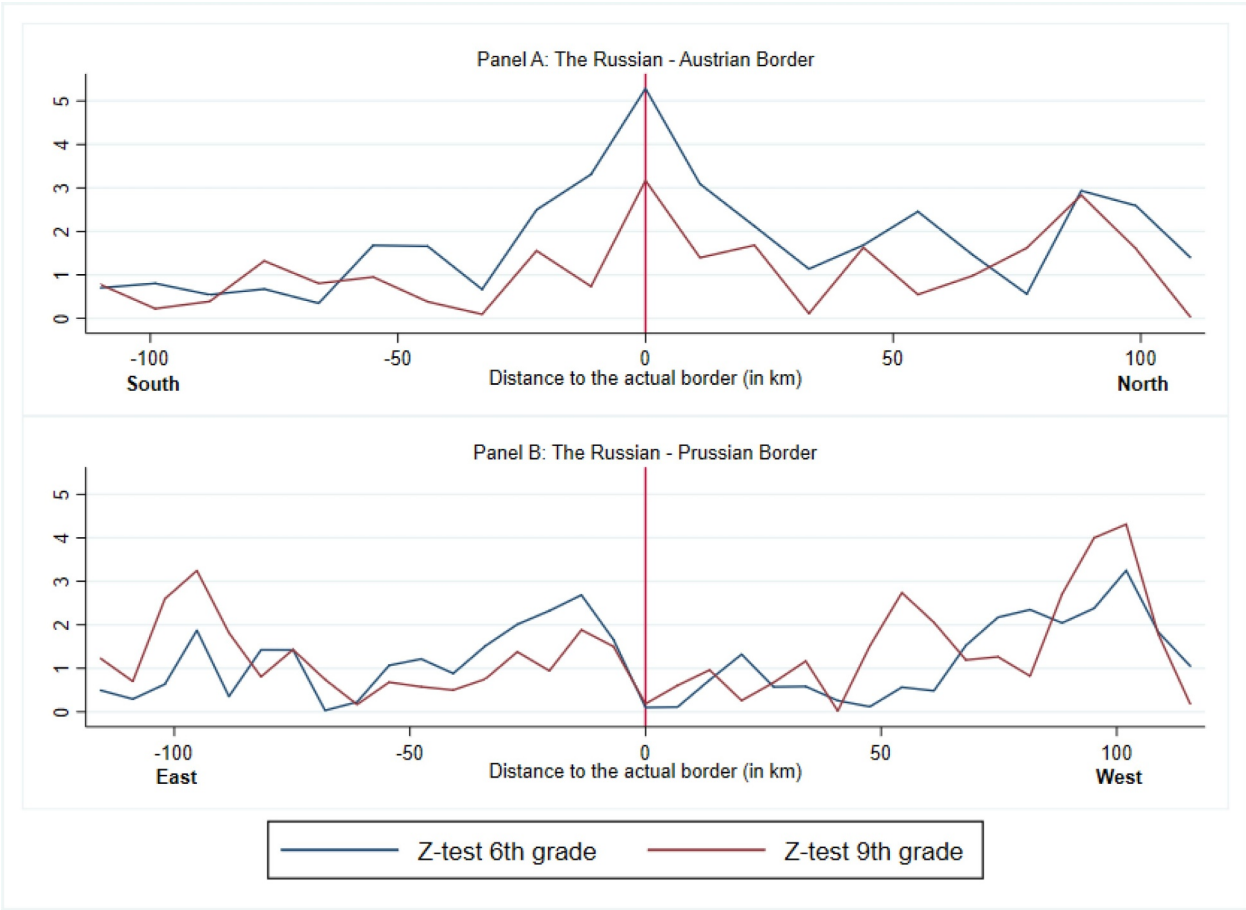

Fig. 9. The placebo experiments. Note: In Panel A, I artificially move step-by-step the Austrian-Russian border by 5km to the north and to the south (negative distance). In Panel B, I artificially move step-by-step the Prussian-Russian border by 5km to the west and to the east (negative distance). For each placebo border, I calculate the Z-test (ratio of a coefficient and a corresponding robust standard errors) for the placebo Partitions dummy coefficient from the baseline specification of Eq. 1. Source: author's computation using the Central Board of Examination data.

4. Channels of persistence

In this section I investigate the channels of persistence. There are many possible explanations of the observed pattern in exam scores. I highlight the social norms channel, which other studies also suggest to be the most important.[34] In particular, the existing literature on the long-lasting effects of the Central and Eastern European Empires underlines the importance of the 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. (2016) and Karaja (2013) show that the government polices introduced by the Hapsburg and Ottoman Empires still affect trust toward local state and acceptance of corruption. My results show that people living in the former Austrian Empire have a more positive social norm with respect to their *local education system*, even though they do not necessarily have a more positive general social norm toward intelligence and higher (non-local) education. I also discuss alternative channels, in particular, skill-biased migration and urbanization patterns.

4.1. Social norms toward local education system

I provide four pieces of empirical evidence to argue that the Partitions of Poland have created different social norms toward the local education system. Although neither of them would be persuasive in isolation, when taken together they suggest that social norm is an important transmission mechanism of past institutions. Firstly, I develop a simple, norm-based model of student schooling efforts, which is consistent with my finding that the effect of the Austrian Empire is larger in the case of the low-stakes exam then in the case of the high-stakes one. Secondly, I use the data on proxies for social values to directly compare people's attitudes toward

[34] See, for instance, Putnam et al. (1994), Akerlof and Kranton (2010), Cassar et al. (2014), Sakalli (2018), and Feir (2016).

education across the former Partition borders. Thirdly, 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. Finally, consistent with the general social-norm channel, I provide evidence for the strong effect of the Partitions on political outcomes, namely voter turnout and pro-EU votes.

4.1.1. Low-stakes vs. high-stakes 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 all input into education) of other grade-invariant characteristics X_i and of an idiosyncratic shock ϵ_{ig}. Then:

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

I follow (Akerlof and Kranton, 2002) and model the level of schooling effort as a function of standard pecuniary incentives (such as the direct costs and benefits of education) and social-norm based incentives, penalizing the individual for not coping 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 effort depends on economic forces and social expectations.

Next, consider a minor modification of the student utility function outlined above. 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-stakes 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 the 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-stakes exam score, the maximization problem is simpler. This exam score does not matter for the future education career, therefore it will not have an impact on the future wages. The utility function 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,[35] I conclude that:

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

This is consistent with the empirical results presented in Section 3, namely that the effect of the Austrian Empire is larger in the case of low-stakes exams, than in the case of high-stakes exams (Table 5). 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 check 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 (Czapiński and Panek, 2013), which include data from over 45,000 individuals from almost all counties in Poland. The data does not offer direct measures of social norms with respect to local schools. Nevertheless, the questionnaire asks whether education is important for a good life, whether respondents are satisfied with the education they have received and their desired level of education

[35] This is a stronger assumption than in the case of the regression discontinuity design.

for their children. The variables are described in Table A3. A second source, with a smaller sample size, is the two waves (2006 and 2010) of 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 the priority of education in governmental spending, private expenditure on education and opinions on the role of intelligence and skills in life success. A third source is the first wave (2010) of the Education Value Added survey (Dolata et al., 2013; 2015) (EVA, conducted by the Education Research Institute in Warsaw), which includes approximately 10,000 parents and asks about the role of family tradition in school selection. For detailed descriptions of the variables from the LiTS and EVA, see Table A2.

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

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

where notation is as in Eq. (1). Additionally, I control for a set of county-level socio-economic characteristics X_c, which are described in Table A1. I pool the two waves of the survey (cross-sections) and include wave fixed effects κ_w. Since there are fewer counties (NUTS4) than municipalities (NUTS5), and to ensure enough statistical power, I use the total sample, and expand the bandwidth to 100km. I employ a one -dimensional polynomial in distance in order to avoid an over-fit of the model. Depending on the outcome variable, I use Probit or Ordered Logit estimators.

In the case of the LiTS and EVA surveys, the samples are based on only several PSUs that are located near the former borders of the Partitions. I use Eq. 3, with the whole sample and 100km bandwidth, but in this case, I do not include a polynomial in location or distance.[36] Beyond this limitation, one should 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 presents estimates of the effect of the Partitions of Poland using the sample from Social Diagnosis and the reported numbers are the marginal effects at the border. Table 8 reports the average marginal treatment effects from the LiTS and EVA samples. The results show that people living in the former Austrian partition are around 5pp. more likely to say that education is important for a good life (Table 7, Columns 1 and 2) and around 24pp. 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 5pp. more likely to claim that family tradition is important in their local school choice (Table 8, Columns 7 and 8). However, at the same time, they are around 20pp. *less* likely to desire higher education for their children (Table 7, Columns 5 and 6) and 22pp. *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 with the education they received (Table 7, Columns 3 and 4), but spend less on the education of their children (Table 8, Columns 5 and 6).

These results can be interpreted as evidence for a positive social norm toward *local* education institutions. People from the former Austrian Empire perceive education as important in their life, want more funds directed to public schools[37] and underscore the long-run relationships of their families with local schools.[38] On the other hand, their attachment to local schools might lead to a lower trust in non-local ones. This would explain why people desire less to send their children to an institution of higher education, which is usually located outside their local communities. Nevertheless, they also perceive intelligence and skills as less important and, even though they might look at these outside the education context, this is clearly a puzzling result. Overall, in 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 schools, that is $E(Austria) > E(Russia)$, but lower social norms 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 or 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 the share of children aged 3–5 who attend an institution of pre-education. I use the regressions specified as in Eq. (1).[39] The results show that the kindergarten attendance ratio is higher on the Austrian side by 3–8 pp compared to the Russian side. On the other hand, there is no significant difference

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

[37] The stronger belief about the priority of education in governmental spending in the Austrian partition can be alternatively explained by the poor quality of local public education. However, if this is the case we would rather observe a negative impact of the Austrian Empire on the exam scores. The results show the opposite. Moreover, there is no systematic difference between the Austrian and Russian partition in terms of 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).

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

[39] Variables are at the municipality level. In all the regressions I use 50 km bandwidth and quadratic polynomial in longitude and latitude. I pool years and use a random effect estimator.

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. (3 = no)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
The Russian - Austrian border								
Partitions effect	−.047	−.054	−.031	−.065	.176	0.22	0.053	0.133
(Russia = 1)	(.022)*	(.021)**	(.046)	(.044)	(.083)*	(.072)**	(.008)***	(.04)***
Mean of outcome	0.06	0.06	0.27	0.27	0.81	0.81	2.4	2.4
R^2	0.16	0.16	0.04	0.04	0.06	0.07	0.02	0.02
Observations	8364	8364	8259	8259	2496	2496	8351	8351
The Russian - Prussian border								
Partitions effect	.038	0.033	−.052	−.074	0.038	.083	−.128	−.102
(Russia = 1)	(.016)*	(.017)*	(.059)	(.057)	(.086)	(.087)	(.035)***	(.038)**
Mean of outcome	0.06	0.06	0.25	0.25	0.75	0.75	2.3	2.3
R^2	.15	0.15	0.02	0.02	0.08	0.091	0.01	0.02
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 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%. 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 dependent variables are described in Table A3. The 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)
The Russian - Austrian border								
Partitions effect	−.085	−.24	.08	0.22	.184	0.683	−.051	−.053
(Russia = 1)	(.082)	(.128)+	(.066)	(.062)***	(.224)	(.313)*	(.032)+	(.053)
Mean of outcome	0.31	0.31	0.32	0.32	-	-	0.24	0.24
R^2	0.06	0.06	0.04	0.09	0.16	0.25	0.01	0.03
Observations	602	602	602	587	233	233	802	802
The Russian - Prussian border								
Partitions effect	.08	−.08	−.07	−.378	0.23	-1.24	0.029	−.015
(Russia = 1)	(.05)	(.137)	(.067)	(.201)+	(.265)	(1.85)	(.026)	(.031)
Mean of outcome	0.31	0.31	0.35	0.35	−	-	0.22	0.22
R^2	.08	0.1	0.03	0.1	0.3	0.37	0.01	0.05
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	No	No
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	Probit	Probit
Source	LiTS	LiTS	LiTS	LiTS	LiTS	LiTS	EVA	EVA

Notes: Robust and clustered standard errors (at 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%. 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 consists of randomly selected PSU, the regressions does not include any polynomial of location or distance. The dependent variables are described in Table A2. The control variables are listed and explained in Table A1. Geographic controls are not included in the regressions based on EVA (Columns 7 and 8) because of the PSUs' anonymity. All the regressions use 100 km bandwidth and include wave fixed effects.

on the Prussian-Russian border.

This 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 or institutions from the 19th century. Unfortunately, no data exist to test this hypothesis. Nevertheless, if this

were 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 parents' willingness to send their children to kindergartens. The higher social norm toward a local education system, including local kindergartens, can thus explain the positive effect of the Austrian Empire on kindergarten attendance.

4.1.4. Political outcomes

The final piece of evidence does not directly look at education, but explores, in general, whether the Partitions can have persisted through social norms. Political outcomes are particularly useful, as they are known to be affected not only by economic factors, but also underlying norms and beliefs (Greenstein, 1969; Moon, 1992; Denny and Doyle, 2008; Graham et al., 2009). In particular, I focus on the effect of the Partitions on voter turnout and votes in favour of EU accession.[40] The former is a proxy of active citizenship, which is likely to be high if people believe that democracy and political institutions matter. Support for EU accession is likely to be affected by positive social norms towards "external" powers. How could the Partitions of Poland affect these norms and beliefs? The Poles living in the Prussian and Austrian Empires could participate in elections and choose their representatives in the national assemblies. Thirteen elections took place in Prussia between 1871 and 1912, and eleven in Austria between 1849 and 1911. On these occasions, Polish national leaders and circles wanted to politically activate the Poles by portraying voting as an act of patriotism (Wojciak, 1981; Trzeciakowski, 2003; Ilski, 2008). In contrast, in the Russian Empire only two elections took place, in 1906–1907, and 1912, providing a limited chance for Poles to fight for autonomy through politics. These suggest that the longer tradition of electoral activism on the former Prussian and Austrian territories, compared to the former Russian, might have a positive effect on willingness to vote. On the other hand, the hostility of the Prussian and Russian Empires towards Poles and the relative tolerance of the Austrian Empire suggest that living in the latter might have a positive effect on inhabitants' acceptance of the "rules of Brussels" and thus on pro-EU votes.

I test these hypotheses by estimating the baseline model described in Eq. (1), but where outcome variables are turnout in presidential (2000, 2005, 2010), parliamentary (2015) and local (1998, 2002, 2006, 2010, 2014) elections, as well as the turnout and results of the EU referendum in 2003. Since urbanization might affect political behaviour, I include controls for population size and density. The data comes from the Polish National Electoral Commission, and Table 10 presents the results. The Russian Partition has a negative, strong and significant effect on voter turnout for the presidential and parliamentary elections, as well as the EU referendum. Crossing the Austrian (Prussian)-Russian border leads to a lower turnout of around 4 (2.75) percentage points for presidential elections, 3.6 (3.2) percentage points for the parliamentary election and 6.8 (5) percentage points for the EU referendum. These results are consistent with Grosfeld and Zhuravskaya (2015). Interestingly, turnout in the local elections in the former Russian partition is 2.85pp higher than in the Austrian partition and around 3.5pp higher than in the Prussian partition. As for the results of the EU referendum, those living in the former Austrian part are more likely to cast pro-EU votes by almost 10pp, compared to the former Russian partition (baseline is 61–62%). The effect is strongly significant. On the other hand, there is no difference in pro-EU voting on the Prussian-Russian border.

The higher turnout in the central-governance elections (presidential, parliamentary and referendum) in the Austrian and Prussian partitions, and the lower turnout in the local-government elections, is partially consistent with the social-norm hypothesis. In the 19th century, Poles were governed by three different central powers, but today they share same president and parliament and thus differences in voter turnout might not arise from different qualities of government, but rather from heterogeneous social norms towards political participation. The rule of law and frequent elections in the Austrian and Prussian Empires created positive norms and trust towards central democratic institutions (Becker et al., 2016) and people perceive the central level as more important than the local. On the other hand, the absolutist and undemocratic rules of the Russian Empire shifted people's attention towards local governments. An alternative explanation could be that local governments are somehow different in the former Russian partition leading to different political participation. But this would imply that they are either more efficient (so people perceive them as important) or less inefficient (so people want to replace them). More research is needed to fully explore this possibility.

The positive effect of the Austrian Empire on the number of pro-EU votes is fully consistent with the social-norm hypothesis. As discussed in Section 2, the Hapsburgs were relatively tolerant and allow Poles much autonomy, which might have engendered trust in "external" forces and raised the acceptance of the EU.[41] This finding thus points to the interaction between identity and institutions as an important condition for the long-run persistence of institutions. The next chapter discusses this implication in more depth.

4.2. Migration

Modern skill-biased migration might be an alternative explanation of the observed positive effect of the Austrian Partition. If, for instance, only high achievers migrate to the former Austrian partition and only low achievers to the former Russian partition, one

[40] For more about the effect of the Partitions of Poland on political outcomes see Grosfeld and Zhuravskaya (2015), Kowalski (2000), and Zarycki (2007).

[41] An alternative explanation based on norms is that the higher pro-EU votes are an outcome of the higher quality of human capital (performance of students) in the former Austrian partition. However, including student performance from 2005 (not available for 2003) into the regression did not change the effect.

should expect to find a significant gap in the average student performance.[42] In order to evaluate this possibility, I follow (Dell, 2010) and exploit the student-level data on 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). 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 Eq. (1)). By construction, 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.[43] The results thus show that modern migration itself is unlikely to explain the observed effect of the Partitions.

Nevertheless, the effect might be driven by historical migration if the selection of migrants was not orthogonal to the current performance of students. My main results could be, for instance, explained by the migration of high-skilled people from the Russian to Austrian parts (or vice versa for the low-skilled) and/or migration of low-skilled people from the Austrian part to third countries. Existing qualitative evidence suggests that the first possibility is unlikely. Labuda (1971) argued that the majority of migrants between the partitions had been seasonal workers, who did not settle permanently.[44] Moreover, the economic situation in Galicia was the hardest and the level of industrialization the lowest of the areas in question, 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 19th century emigrants' skills from partitioned Poland. Zubrzycki (1953) hypothesizes that migrants from the Grand Duchy of Poznań were more likely to be educated than migrants from Galicia and the Congress Kingdom.[45] However, as reported in Abramitzky et al. (2014), late 19th century U.S. immigrants from Austria and Prussia had wages similar to the US population, even though there was substantial variation in wages between 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. In Table 11, Panel A presents the migration balance and Panel B presents migration as a share of the 1910 population for each partition. The numbers show a large outflow of population from the Austrian and Prussian parts. Importantly, however, the Austrian partition does not seem to be unique and, in fact, migration was larger in the Prussian partition. A similar picture emerges from the data on general population characteristics (Panels C and D). The population trends and age structures are similar across the partitions and the Austrian part does not show any migration-related anomalies, for instance, the over or under representation of the middle-aged population group.

The post-World War II forced displacement of Poles may have erased the effect of the Prussian Empire. This migration was from the eastern parts of interwar Poland to the newly joined post-German areas in the west and north of modern Poland (see Fig. 1). At the same time, Germans were expelled to Germany, leaving their houses and agriculture holdings behind. If the areas on the Prussian side of the Prussian-Russian border had had a significant German minority before 1939, then the incoming Polish population may have been more likely to settle there. I argue that this scenario is unlikely. In my analysis, I exclude Silesia and Eastern Prussia, which had historically significant German populations and focus only on the areas with Polish majorities. In order to investigate the presence of German minorities in the borderlands of interest, I use the 1931 Polish census data at the county level on the share of German speakers (GUS, 1938).[46] The upper part of Fig. 10 presents the geographic distribution of German speakers and the shaded areas mark the 50km bandwidth around the former partitions' borders. Table A9 presents the county-level average share of German speakers for the same bandwidth. Even though Germans were concentrated in the former Prussian partition, there is no visible discontinuity at the border with the Russian partition. The average share is 7.17% for the former Prussian borderlands and 6.28% for the former Russian, and the difference is not significant. If anything, the share of Poles is lower on the Russian side. This is because the Congress Kingdom had relatively a high share of Jews (see the lower part of Fig. 10 and Table A9).

4.3. Other channels

The Partitions of Poland may have persisted because of other differences between the Empires, such as urbanization, 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 in Eq. (1),[47] but with the dependent variables transformed to natural logarithms. Similarly to the exam scores, these variables are endogenous, and might reflect the effect of the historical education systems or other channels.

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

[43] 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.

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

[45] Nevertheless, given the size of migration (see Table 11), the majority of migrants were likely uneducated.

[46] Unfortunately, to my best knowledge, there is no census-level data available on the share of displaced people. The census of 1950 provides voivodeship-level data. However, the voivodeships on the Prussian side of the Prussian-Russian border include the areas joined to Poland after WWII and thus the average shares are obviously higher there.

[47] 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 a random effect estimator, for the time-invariant variables I use OLS.

Table 9
Kindergarten attendance.

Dep. Variable:	Kindergarten attendance					
	(1)	(2)	(3)	(4)	(5)	(6)
The Russian - Austrian border						
Partitions effect	−8.45	−5.48	−4.16	−6.28	−2.97	−8.84
(Russia = 1)	(2.33)***	(2.5)*	(3.38)	(2.09)**	(2.29)	(3.21)**
Mean of outcome	47pp	47pp	47pp	51pp	51pp	51pp
R^2	.21	0.21	0.4	0.34	0.35	0.5
Municipalities	301	301	301	373	373	373
Mun. × Time	2107	2107	2101	2611	2611	2606
The Russian - Prussian border						
Partitions effect	−3.81	−3.56	1.43	−2.75	-2.78	−2.89
(Russia = 1)	(3.06)	(3.05)	(3.53)	(2.53)	(2.53)	(3.13)
Mean of outcome	48pp	48pp	48pp	52pp	52pp	52pp
R^2	.36	0.36	0.54	0.41	0.41	0.55
Municipalities	206	206	206	302	302	302
Mun. × 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 standard errors (at municipality level) 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 the 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). The control variables are listed and explained in Table A1. All the regressions use 50 km bandwidth.

Table 10
Political outcomes.

Dep. Variable:	Presidential elections - Turnout		Parliamentary elections - Turnout		Local elections - Turnout		EU referendum - Turnout		EU referendum - Votes "yes"	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
The Russian - Austrian border										
Partitions effect	-3.17	−3.71	−2.67	−3.18	1.96	2.8	−5.76	−6.47	−8.89	−9.31
(Russia = 1)	(.99)***	(.87)***	(.96)**	(.88)***	(1.12)	(.963)**	(.99)***	(.93)***	(1.74)***	(1.56)***
Mean of outcome	52	53	47	47	53	53	51	52	62	61
R^2	0.59	0.60	0.61	0.58	0.22	0.21	0.55	0.57	0.51	0.5
Observations	903	1119	301	373	1505	1865	301	373	301	373
The Russian - Prussian border										
Partitions effect	−2.93	-2.35	−3.82	−2.35	3.78	2.38	−5.25	−4.73	−.07	−.12
(Russia = 1)	(.87)**	(.72)***	(1.06)***	(.92)*	(1.7)*	(1.39)	(1.04)***	(.88)***	(1.96)	(1.7)
Mean of outcome	49	50	41	43	53	52	50	52	61	65
R^2	.81	0.80	0.48	0.46	0.27	0.28	0.59	0.62	0.26	0.44
Observations	618	906	206	302	1030	1510	206	302	206	302
Geographic controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Sample	Rural	All	Rural	All	Rural	All	Rural	All	Rural	All

Notes: Robust and clustered standard errors (at municipality level) are reported in the parentheses. *** denotes significance at the 0,1% level, ** at the 1% level, * at the 5% and + at the 10%. 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 or Prussian territories. In addition, the regressions include a quadratic polynomial of distance and a set of geographic covariates, population size and density. The dependent variables are explained in Table A1. All the regressions use 50 km bandwidth.

Table 6, Column 1 reports estimates of the effect (semi-elasticity) of the Russian Empire on the Prussian-Russian border, while Column 2 reports the effect on the Austrian-Russian border; therefore, changing the sign yields either the effect of the Prussian or Austrian Empire. 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. The Austrian Empire positively affects the demographic characteristics, namely the level of population, density, and migration balance. Different urbanization patterns could be an alternative channel if they affect the quality of education. One mechanism could be through class size (Angrist and Lavy, 1999).[48]

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

Table 11
Historical migration and demographic characteristics.

Year/Partition:	Russian	Austrian	Prussian
Panel A: Migration balance (in thousand)			
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 the 1871-1910s' as % of the 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 groups 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 12
The 19th century education attainment and the modern exam scores.

Dep. Variable:	6th low-stakes exam				9th high-stakes 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 × Attainment	0.282	0.17	.24	.162	0.172	0.11	0.128	0.069
	(.114)*	(.085)*	(.097)*	(.07)*	(.097)+	(.089)	(.095)	(.076)
Mean of outcome	-.016	−.016	0.067	0.067	0.005	0.005	0.053	0.053
R^2	0.32	0.38	0.25	.36	.2	0.22	0.17	0.23
Municipalities × 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	Rural	All	All	Rural	Rural	All	All

Notes: Robust and clustered standard errors (at county or deanery levels) 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-stakes exam score; Columns 5 to 8 the mathematics and science 9th grade high-stakes exam score. Table presents estimates of the effect of the 19th century education 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.

However, as shown in Table 6, there is no difference in class size between the former Austrian and Russian lands. Another mechanism could be school network density. A higher density might lead to higher school competition, which is generally believed to have a positive effect on student performance (Urquiola, 2016). Again, there is no significant difference between the partitions in terms of the number of elementary and secondary schools per squared km.

More broadly, if urbanization were the main channel of persistence, I should find an insignificant and small effect of the Austrian Partition after controlling for urbanization proxies. In regressions reported in Section 3.5, I include continuous variables for school density, population size and density. For the joint rural and urban sample, I additionally include a categorical variable describing the type of municipality.[49] As reported in Table A8, the positive and strong effect of the Austrian Partition is insensitive to the inclusion of demographic characteristics.

Nevertheless, my results may be driven not by the size of urbanization, but rather its quality (e.g. types of settlements, "urban culture"). If formerly Austrian urban areas are somehow different from formerly Russian ones, and this difference affects student performance, I should find a stronger partition effect in more urbanized areas. To check this possibility, I modify the baseline regressions and add an interaction term between the partition dummy and either population size or density. Table A10 shows that

[49] There are four categories of municipalities in Poland: rural, urban-rural, urban and cities.

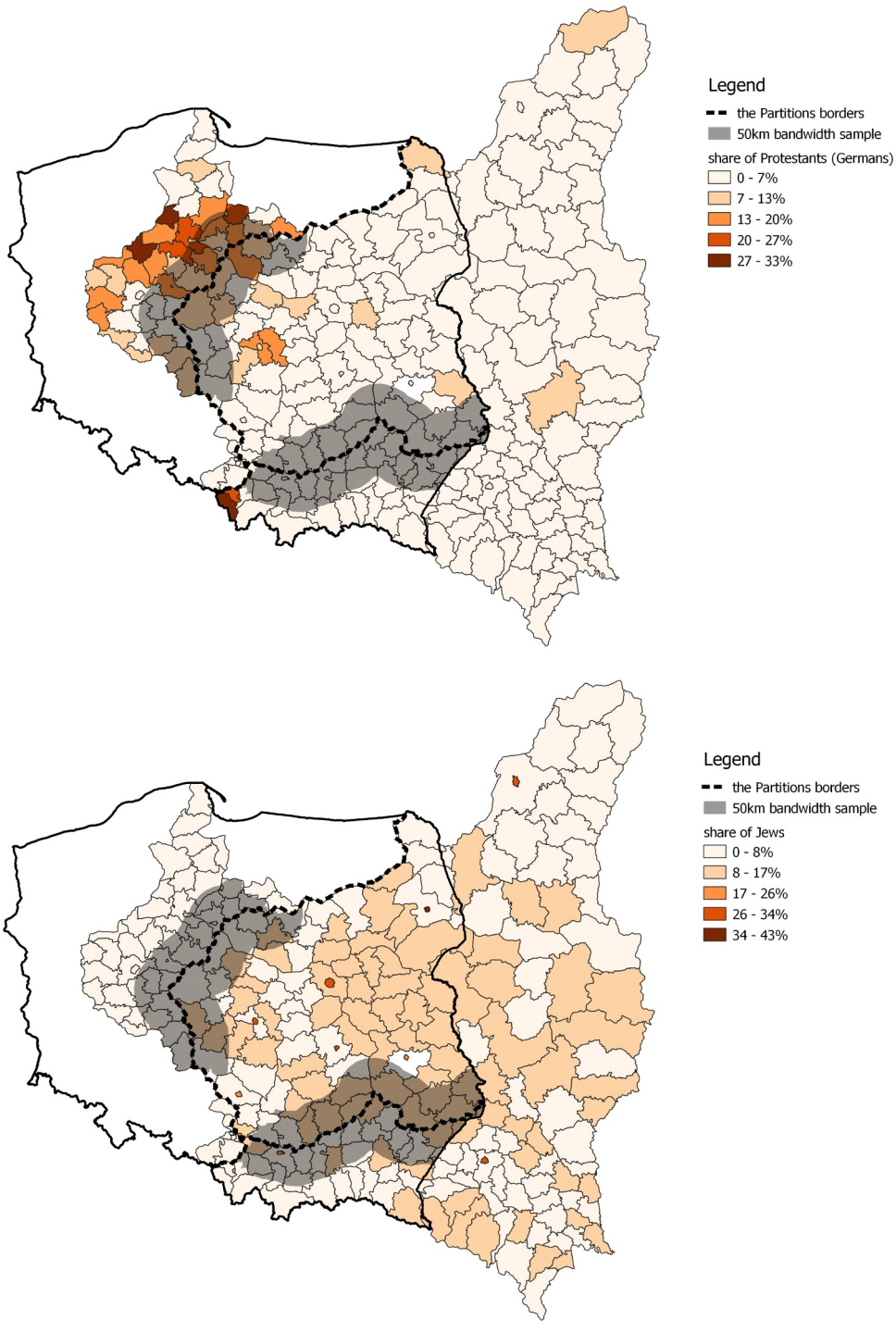

Fig. 10. The share of Protestants (Germans) and Jews in 1931. Note: the map shows the county-level share of Protestants and Jews. The borders of the Partitions of Poland are marked by the dashed line. The modern borders of Poland are marked by the solid line. Source: author's computation using the Census of 1931. Own visualization based on GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

neither of the interaction terms is statistically significant, meaning that there is no evidence for a heterogeneous effect by the urbanization measures. In addition, I run the baseline regressions separately for the areas with population density larger and smaller than the median density. The results are virtually the same (available upon request).

Finally, as Table 6 shows, 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

driving mechanisms. Finally, there is a significant positive effect on the share of people with higher education. This is consistent with the social-norm hypothesis, as one would expect that social norms toward education influence not only the performance of students, but also other education 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 arguments that identity is associated with social norms affecting an individual's schooling choices, the school-student relationship and student achievements. A similar hypothesis is also explored by Sakalli (2018), who documents that the Muslim identity of Turks has been reinforced by past coexistence with Armenians, which in turn, has changed the long-run social and cultural norms toward secular education.

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

5.1. Conceptualization

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

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

Utility comes from the 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, which are used to eradicate individuals' ethnic identity, will impose higher costs on them. In the case of the Partitions of Poland, the Prussian education system required Polish students to learn in German and to study an anti-Polish curriculum. The cost of education included then an additional cost of learning a foreign language and the intrinsic discomfort coming from ethnic intolerance. Conversely, the Austrian system offered similar institutions, but with Polish as the language of instruction and without anti-Polish content. 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 contributed to the higher schooling effort.

Next, suppose that a social norm about some behaviour emerges when all individuals consistently find this behaviour as optimal. This social norm is then transmitted through generations and affects the schooling effort in the second period (Bisin and Verdier, 2001; Patacchini and Zenou, 2011; Spolaore and Wacziarg, 2016). The difference for future generations is that past institutions have affected individuals' optimal behaviour not through the standard part of the utility function, but through the social norm part. Therefore, in the second period, individuals also face a social norm with respect to schooling effort. Using Eq. (2) and assuming, for simplicity, that the wage premium and cost parameters in the second period are equal to unity ($w_2 = a_2 = 1$), the second period utility function is:

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

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

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

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

This simple model implies that the relatively lower costs of education in the Austrian Empire became a crucial factor for the formation of this social norm and thus for future schooling efforts.

5.2. Evidence

The analysis so far has 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 education attainment,

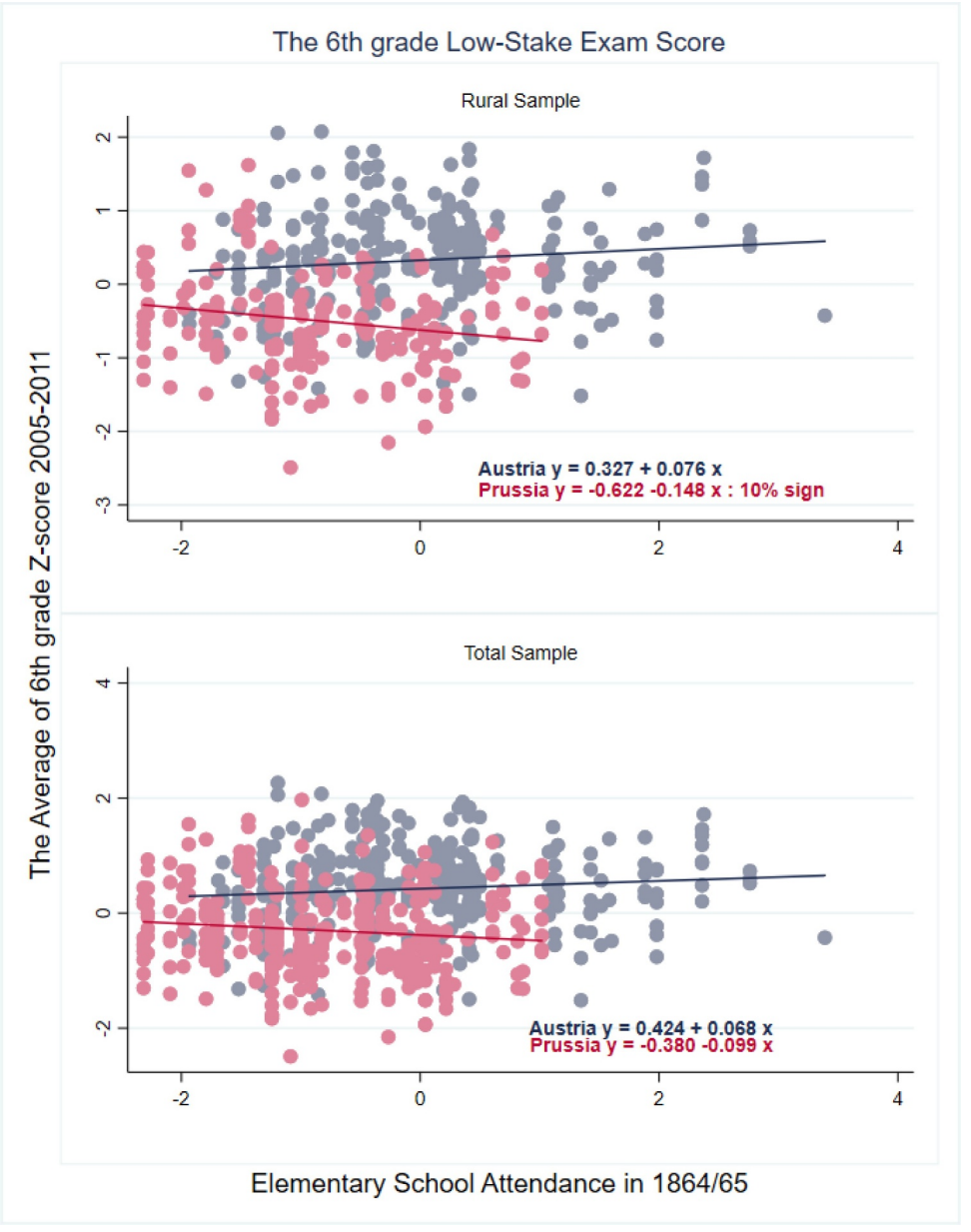

Fig. 11. Historical education attainment and the modern performance of students. Note: the figures present the standardized elementary school education attainment in 1864/65 (x-axis) plotted against the municipality average (2005–2011) of the standardized 6th grade low-stakes exam (y-axis). The former Austrian partition is in grey, the former Prussian partition is in red. The line is a fitted line from a regression of the 6th grade exam score on the education 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. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to measure the *within*-partition variation in the exposure to 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 past education attainment, as measured by the elementary school enrolment ratio, on the modern-day quality of education in the former Austrian Empire, but a null or negative effect in the former Prussian Empire.

The historical data on education characteristics comes from the 19th century censuses. In the case of the Prussian Empire, the source is the Ifo Prussian Economic History Database (Becker et al., 2014). 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 assign 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 were administration units of

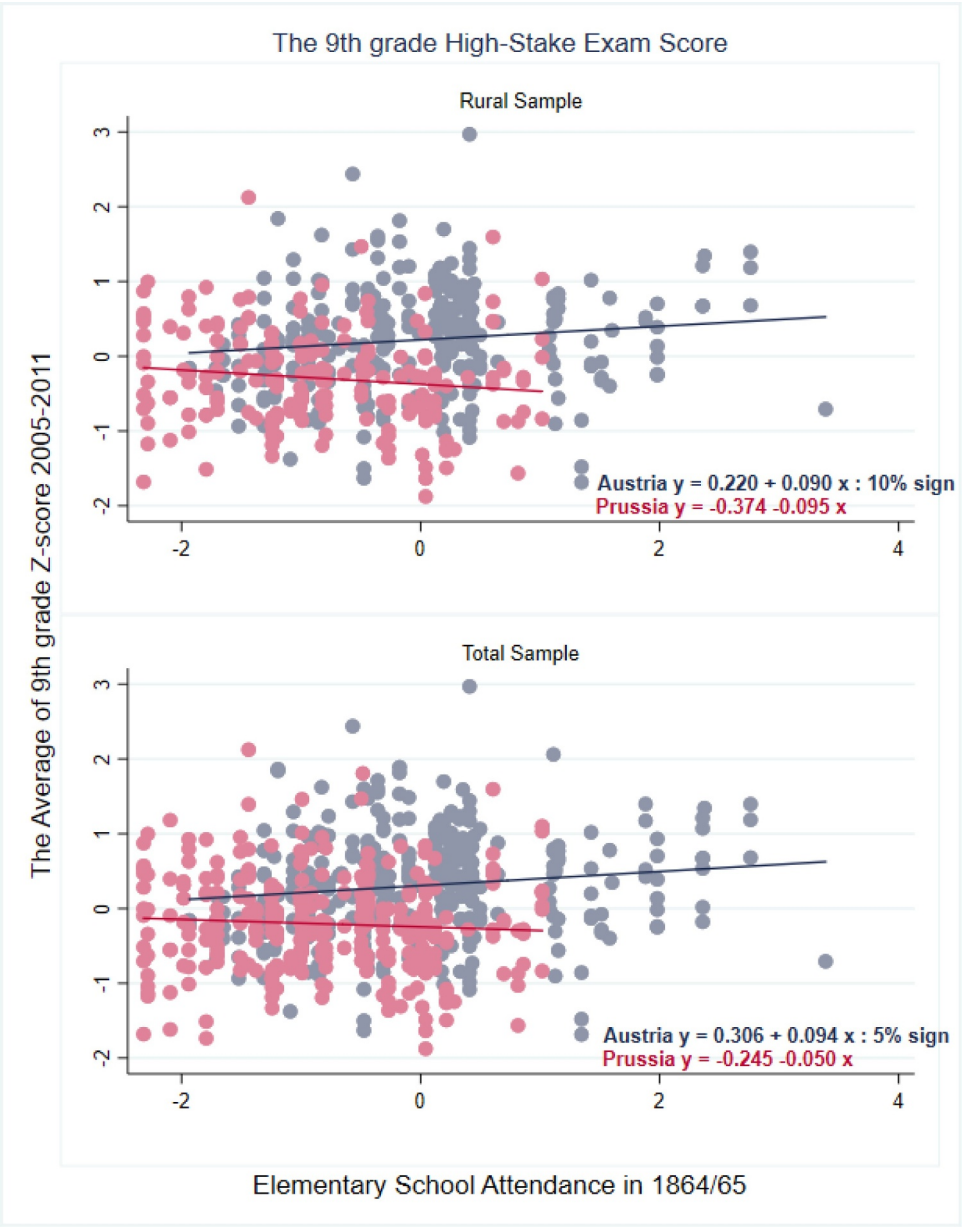

Fig. 12. Historical education attainment and the modern performance of students. Note: the figures present the standardized elementary school education attainment in 1864/65 (x-axis) plotted against the municipality average (2005–2011) of the standardized 9th grade high-stakes exam (y-axis). The former Austrian partition is in grey, the former Prussian partition is in red. The line is a fitted line from a regression of the 9th grade exam score on the education 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. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the Catholic Church. As there is no GIS map of deaneries from the Austrian Empire, I manually match modern municipalities with their historical deaneries using information from Dobrowolski (1886) and the geographic dictionary by Sulimierski et al. (1895). Unfortunately, for the Austrian part, only the census of 1865 offers data dis-aggregated to the deanery level. In order to keep a comparable time frame, I use only data from the Prussian census of 1864. The variable of interest is the total education 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 censuses, I standardized the measures, so that for each partition they have a mean of zero and a standard deviation of one.

I first document a simple correlation between the standardized 19th century education attainment and the average of the modern exam scores (2005–2011). Fig. 11 presents the correlation for the 6th grade low-stakes exam score and Fig. 12 for the 9th grade high-

stakes exam, in a breakdown by the rural and total samples. Consistently with the hypothesis, we 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 are likely to be biased. For instance, the favourable location of a municipality might affect its long-run prosperity and influence historical education attainment along with the current performance of students. In order to limit bias, I run regressions of the standardized 6th or 9th grade exam scores on historical education 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 a dummy for Austria and interact it with the historical measure:

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

where y_{mcwt} is the outcome variable for a municipality m from county (deanery) c and partition p at time t (available for 2005–2011). A_{cp} is the education attainment from the 1860s, which is available at the county (deanery) level, Aus_p is a dummy for the former Austrian partition and 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 a random effect estimator. The standard errors are clustered at the county (deanery) level.

Table 11presents the estimates of Eq. (6). Columns 1 and 2 show that the correlations between the education attainment and the 6th grade exam scores are significant and negative for the former Prussian partition. One standard deviation increase in the attainment is connected with a 0.14 - 0.26σ decrease in the student performance. Importantly, these correlations are 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 would have a 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. The social norms affecting student performance were more likely to emerge in municipalities with a larger attachment to the historical Austrian education system. Alternatively, because of the *negative* interaction between the institutional quality and identity, the more intensive historical exposure to Prussian education led to a stronger opposing social norm toward the education system, lowering schooling effort. Furthermore, using arguments from Section 4.1.1, the weaker relationship in the case of the 9th grade high-stakes exam points to the importance of social norms as a channel of persistence. Nevertheless, the inclusiveness of the Austrian *secondary* education or the existence of two universities and one technical college in Galicia[50] might be another source of persistence.

6. Conclusions

This paper has argued that the Partitions of Poland provide a unique natural experiment for studying the determinants of institutional persistence. First, I exploited this setting to investigate the long-lasting effect of the 19th century education systems, which were imposed by Austria, Prussia, and Russia, on modern education outcomes. Despite the modern similarities of the former borderlands of the Empires, I estimated 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 comparable with the Black vs. White achievement gap in the US.

How can we explain these results, and what can we learn about the persistence of institutions? The main hypothesis here posits that an interaction between institutional quality and identity might be crucial for the creation of a positive social norm toward institutions, and thus for long-lasting persistence. In particular, because the Austrian education system actively supported 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 an almost identical education system was used as a tool of Germanisation, and no positive social norm affects the modern performance of students.

This result might be of crucial importance for policymakers who wish to improve the situation of permanently underdeveloped regions. For instance, the implication of my hypothesis is that good education 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 (e.g. for Hungarians in Romania, Poles in Lithuania or Russians in Ukraine) as an example of such a policy. On the other hand, this paper suggests that large interventions aimed at equalizing education differences, as carried out by the post-WWII communist government in Poland, might have had a large effect on the quantity of education, but a limited effect on the quality of education.

The proposed sources of persistence can also shed light on the existing, and often puzzling, findings in the literature on in-

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

stitutional persistence. For instance, a study of Africa, by Michalopoulos and Papaioannou (2013), reported that the pre-colonial ethnic institutions matter for long-run development to a much larger extent than the national ones created by the colonial powers. Similarly, in the Indian context, Iyer (2010) reported a negative effect of British colonial rules and a positive one 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 existing social identities. In contrast, the external powers imposed institutions based on borders which broadly ignored ethnic divisions.

The interaction between institutions and ethnic identity can be perceived as an ingredient of inter-ethnic inequalities and modes of cooperation. As suggested by Jha (2013), limited ethnic assimilation might lead to "the presence of a nonreplicable and non-expropriable source of inter-ethnic 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 inter-ethnic complementarity is a higher specialization in an economy, which, in turn, increases productivity.[51] In the context of Native American tribes, Dippel (2014) presents evidence that the forced integration of linguistically homogeneous sub-tribal groups has had a negative effect on long-run economic development through the lowering of the quality of local governance. I have added 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.[52]

Finally, this paper has contributed to the literature on the formation of human capital by pointing to the importance of social norms for the quality of education. Many authors have claimed 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 analysing institutions outside of their social context might be insufficient.

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 - Controls		
Density:	Population density.	2005-11
Expenditure:	Local government (municipality) total expenditure per capita.	2005-11
Education expenditure:	Local government (municipality) education expenditure per capita.	2005-11
Kindergarten attendance:	Rate of student pre-elementary schools' attendance.	2005-11
Migration:	Migration balance per 1000 inhabitants.	2005-11
Population:	Total population.	2005-11
Population density:	Total population per km^2	2005-11
Secondary school scholarization:	Rate of student secondary schools' attendance.	2005-11
Elementary schools per km^2:	The number of public elementary schools per km^2	2005-11
Secondary schools per km^2:	The number of public secondary schools per km^2	2005-11
Unemployment rate:	Share of unemployed among the active population.	2005-11
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.	2005-11
Education Value Added:	The estimates of the Education Value Added (gain between 6th and 9th grade).	2013
Election turnout:	The voter turnout for presidential, parliamentary, local elections, and the EU-accession referendum.	1998-14
The EU-accession result:	The share of votes in favour ("Yes") of the EU-accession during the referendum.	2003

Notes: All the variables come from the Central Statistical Office of Poland, except the variables for 2009, which come from the System of Education Information, for the education value added, which comes from the Education Value Added Team, for the geographical controls, which come form *WorldClim.org*, and for the election turnout and results, which come from the Polish National Electoral Commission.

[51] On the other hand, because of the heterogeneity of preferences, fragmented societies are more likely to have inefficient and poorly-managed public goods (La Porta et al., 1999). Also, the existence of heterogeneous ethnic groups, which are clearly distinguishable might motivate a 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).

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

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.
Agree 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 home-town:	A categorical variable with a six-degree scale where 1 is "Cities larger than 500 thousand" and 6 is "Villages"

Table A4
The Akaike information criteria.

Border:	Russian-Prussian		Russian-Austrian	
Outcome Exam:	6th grade (1)	9th grade (2)	6th grade (3)	9th grade (4)
None	4612.0893	4750.5143	5339.8517	5137.3178
Linear	4590.8342	4738.2301	5287.7064	5099.3671
Quadratic	**4590.0828**	**4731.9785**	5281.559	5099.2385
Cubic	4591.2627	4733.8847	5262.8031	5095.0966
Quartile	4591.2627	4733.8847	**5262.3302**	**5094.2096**

Notes: The table shows the Akaike Information Critera for the regressions of either 6th or 9th grade exam score on the partition dummy *D*, which equals 1 for the former Russian areas and 0 for either the former Prussian (Columns 1–2) or Austrian (Columns 3–4) territories, and different polynomials of longitude and latitude. Each row represents different polynomial order. The regressions use 50 km bandwidth.

Table A5
Robustness: Polynomials in latitude and longitude.

Dep. Variable:	6th grade low-stakes exam			9th grade high-stakes exam		
Polynomial/Bandwidth:	< 50km (1)	< 75km (2)	< 100km (3)	< 50km (4)	< 75km (5)	< 100km (6)
Panel A : The Russian - Austrian border						
Linear	−.570	−.695	−.631	−.453	−.492	−.408
	(.116)***	(.108)***	(.103)***	(.124)***	(.111)***	(.108)***
Quadratic	−.561	−.622	−.615	−.409	−.432	−.390
	(.125)***	(.115)***	(.110)***	(.131)***	(.117)***	(.113)***
Cubic	−.549	−.576	−.551	−.392	−.407	−.332
	(.123)***	(.115)***	(.110)***	(.133)***	(.120)***	(.118)***
Quartile	-.558	−.566	−.549	−.405	−.389	−.320
	(.123)***	(.117)***	(.110)***	(.132)***	(.122)***	(.117)***
Mean of outcome	0.1	0.09	0.06	0.02	0.03	0
Municipalities × Time	2107	2981	3688	2106	2981	3681
Municipalities	301	426	527	301	426	527
Panel B : The Russian - Prussian border						
Linear	-.030	.165	.036	-.096	.340	.246
	(.149)	(.141)	(.126)	(.164)	(.154)**	(.133)*
Quadratic	−.059	.129	.041	-.132	.317	.244
	(.156)	(.141)	(.127)	(.169)	(.156)**	(.137)*
Cubic	−.060	.099	.044	-.135	.294	.163
	(.154)	(.141)	(.132)	(.170)	(.156)*	(.140)
Quartile	-.101	.033	.049	-.178	.253	.131
	(.153)	(.143)	(.132)	(.170)	(.157)	(.141)
Mean of outcome	−.22	−.23	−.27	−.04	−.05	−.12
Municipalities × Time	1442	2135	2898	1442	2135	2894
Municipalities	206	305	414	206	305	414
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Modern controls	No	No	No	No	No	No
Sample	Rural	Rural	Rural	Rural	Rural	Rural

Notes:Robust and clustered standard errors (at municipality level) 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-stakes exam score; Columns 4 to 6 - the mathematics and science 9th grade high-stakes 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 A6
Robustness: Polynomials in latitude and longitude, the total sample.

Polynomial/Bandwidth:	6th grade low-stakes exam			9th grade high-stakes exam		
Polynomial /	< 50km (1)	< 75km (2)	< 100km (3)	< 50km (4)	< 75km (5)	< 100km (6)
Panel A : The Russian - Austrian border						
Linear	-.614	−.713	−.664	−.457	−.494	−.445
	(.106)***	(.098)***	(.093)***	(.110)***	(.099)***	(.095)***
Quadratic	−.554	−.618	−.635	−.402	−.431	−.433
	(.116)***	(.105)***	(.099)***	(.122)***	(.107)***	(.103)***
Cubic	−.533	−.569	−.577	−.384	−.411	−.393
	(.116)***	(.107)***	(.101)***	(.125)***	(.111)***	(.107)***
Quartile	-.547	−.556	−.573	−.400	−.391	−.376
	(.116)***	(.108)***	(.101)***	(.123)***	(.113)***	(.107)***
Mean of outcome	0.19	0.19	0.16	0.1	0.1	0.08
Municipalities × Time	2606	3640	4508	2605	3641	4502
Municipalities	373	521	645	373	521	645
Panel B : The Russian - Prussian border						
Linear	-.134	.040	-.049	−.013	.243	.173
	(.121)	(.114)	(.100)	(.123)	(.118)**	(.100)*
Quadratic	−.153	.006	-.045	−.090	.211	.181
	(.128)	(.115)	(.101)	(.132)	(.120)*	(.103)*
Cubic	−.151	−.025	−.055	−.092	.195	.130
	(.128)	(.115)	(.104)	(.132)	(.120)	(.106)
Quartile	-.191	−.091	−.046	−.129	.147	.106
	(.128)	(.116)	(.105)	(.132)	(.120)	(.106)
Mean of outcome	−.15	−.15	−.18	−.03	−.02	−.08
Municipalities × Time	2114	3094	4214	2114	3094	4210
Municipalities	302	442	602	302	442	602
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Modern controls	No	No	No	No	No	No
Sample	All	All	All	All	All	All

Notes: Robust and clustered standard errors (at municipality level) 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-stakes exam score; Columns 4 to 6 - the mathematics and science 9th grade high-stakes 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.

Table A7
Robustness: Polynomials in distance.

Dep. Variable:	6th grade low-stakes exam			9th grade high-stakes exam		
Polynomial/Bandwidth:	< 50km (1)	< 75km (2)	< 100km (3)	< 50km (4)	< 75km (5)	< 100km (6)
Panel A : The Russian - Austrian border						
Linear	−.575	−.671	−.577	−.468	−.477	−.326
	(.170)***	(.141)***	(.133)***	(.149)***	(.125)***	(.124)***
Quadratic	−.483	−.434	−.556	−.273	−.361	−.459
	(.235)**	(.198)**	(.177)***	(.217)	(.177)**	(.160)***
Cubic	−.504	−.424	−.301	−.574	−.224	−.206
	(.343)	(.267)	(.238)	(.300)*	(.235)	(.216)
Quartile	-.393	−.444	−.388	−.470	−.234	−.249
	(.330)	(.265)*	(.235)*	(.290)	(.234)	(.212)
Mean of outcome	0.1	0.09	0.06	0.02	0.03	0
Municipalities × Time	2107	2981	3688	2106	2981	3681
Municipalities	301	426	527	301	426	527
Panel B : The Russian - Prussian border						
Linear	-.068	.107	.001	.020	.431	.292
	(.162)	(.138)	(.120)	(.172)	(.150)***	(.131)**
Quadratic	−.384	−.221	.073	-.118	−.162	.275
	(.240)	(.203)	(.181)	(.249)	(.209)	(.190)
Cubic	−.448	−.492	−.356	−.324	−.344	−.345
	(.321)	(.265)*	(.234)	(.332)	(.272)	(.240)
Quartile	-.816	−.436	−.550	.065	.048	-.341

(*continued on next page*)

Table A7 (*continued*)

Dep. Variable:	6th grade low-stakes exam			9th grade high-stakes exam		
Polynomial/Bandwidth:	< 50km (1)	< 75km (2)	< 100km (3)	< 50km (4)	< 75km (5)	< 100km (6)
	(.460)*	(.339)	(.287)*	(.450)	(.345)	(.297)
Mean of outcome	−.22	−.23	−.27	−.04	−.05	−.12
Municipalities × Time	1442	2135	2898	1442	2135	2894
Municipalities	206	305	414	206	305	414
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Modern controls	No	No	No	No	No	No
Sample	Rural	Rural	Rural	Rural	Rural	Rural

Notes: Robust and clustered standard errors (at municipality level) 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-stakes exam score; Columns 4 to 6 - the mathematics and science 9th grade high-stakes 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 A8
Robustness: Polynomials in latitude and longitude, including socio-economic covariates.

Dep. Variable:	6th grade low-stakes exam			9th grade high-stakes exam		
Polynomial/Bandwidth:	< 50km (1)	< 75km (2)	< 100km (3)	< 50km (4)	< 75km (5)	< 100km (6)
Panel A : The Russian - Austrian border						
Linear	-.388	−.465	−.444	−.296	−.327	−.250
	(.117)***	(.106)***	(.100)***	(.128)**	(.113)***	(.107)**
Quadratic	−.424	−.451	−.472	−.309	−.318	−.271
	(.125)***	(.113)***	(.107)***	(.133)**	(.116)***	(.110)**
Cubic	−.413	−.412	−.416	−.294	−.297	−.218
	(.124)***	(.114)***	(.109)***	(.136)**	(.122)**	(.116)*
Quartile	-.417	−.404	−.417	−.305	−.281	−.210
	(.124)***	(.115)***	(.109)***	(.134)**	(.123)**	(.116)*
Mean of outcome	0.1	0.09	0.06	0.02	0.03	0
Municipalities × Time	2102	2973	3679	2101	2973	3672
Municipalities	301	426	527	301	426	527
Panel B : The Russian - Prussian border						
Linear	-.046	.112	.090	-.180	.294	.264
	(.134)	(.131)	(.121)	(.153)	(.148)**	(.130)**
Quadratic	−.100	.087	.103	-.217	.284	.274
	(.142)	(.132)	(.121)	(.160)	(.151)*	(.133)**
Cubic	−.105	.075	.115	-.220	.269	.203
	(.141)	(.132)	(.125)	(.160)	(.152)*	(.137)
Quartile	-.082	.038	.125	-.209	.234	.165
	(.141)	(.134)	(.126)	(.159)	(.152)	(.137)
Mean of outcome	−.22	−.23	−.27	−.04	−.05	−.12
Municipalities × Time	1442	2135	2898	1442	2135	2894
Municipalities	206	305	414	206	305	414
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Modern controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Rural	Rural	Rural	Rural	Rural	Rural

Notes: Robust and clustered standard errors (at municipality level) 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-stakes exam score; Columns 4 to 6 - the mathematics and science 9th grade high-stakes 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.

Table A9
Ethnic and religious composition in 1931.

Border:	Russian-Prussian					Russian-Austrian				
Variable/Partition:	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)
Panel A : Urban and rural areas < 50km from the borders (in pp.)										
Germans	7.172	4.656	6.28	5.153	0.892	0.255	0.835	0.089	0.109	0.166
Poles	91.031	4.484	86.833	3.72	4.198*	89.806	10.796	89.494	3.988	0.311
Jews	1.804	3.219	6.63	2.771	−4.825**	8.274	4.862	10.736	4.177	−2.462
Greek-Catholics	0.061	0.039	0.038	0.008	0.023	3.958	11.859	0.041	0.019	3.917
Orthodox	0.112	0.147	0.099	0.057	0.013	3.417	9.405	0.097	0.126	3.32
Number of counties	29		6			26		10		
Panel B : Rural areas < 50km from the borders (in pp.)										
Germans	8.631	5.467	7.543	5.481	1.089	0.157	0.406	0.082	0.128	0.075
Poles	90.873	5.249	91.051	5.327	−.178	93.472	11.243	95.499	1.783	−2.027
Jews	0.375	0.869	1.065	0.73	−.69+	3.099	2.244	4.584	1.786	−1.484+
Greek-Catholics	0.05	0.038	0.035	0.014	0.015	4.412	12.999	0.032	0.011	4.38
Orthodox	0.061	0.096	0.057	0.028	0.004	3.89	10.589	0.042	0.041	3.848
Number of counties	24		6			25		9		

Notes: Means and standard deviations for the counties from 1931, located at most 50km either from the former Russian-Prussian or Russian-Austrian borders. The counties located in Ślşskie voivodeships are excluded. Germans (Poles) is a share of German (Polish) speaking people in the total or rural populations. The variables are defined in the same way for Jews, Greek-Catholics and Orthodox, but ethnicity is based on religion affiliation, not language. ** denotes significance at the 1% level, * at the 5% level and + at the 10% level.

Table A10
The interactions between the Partitions effect and urbanization

Dep. Variable:	6th grade low-stakes exam				9th grade high-stakes exam			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: The Russian - Austrian border								
Partitions effect	−.467	−.580	−.604	−.531	−.361	−.311	−.393	−.369
(Russia = 1)	(.149)**	(.199)**	(.126)***	(.119)***	(.158)*	(.218)	(.135)**	(.131)**
Interaction term	.00005	.002	0.00006	−.00003	.0001	.0002	0.00001	−.00007
	(.0001)	(.002)	(.00004)	(.0002)	(.0001)	(.002)	(.00004)	(.0002)
Mean of outcome	.1	0.1	0.19	0.19	0.02	0.02	0.1	0.1
R^2	.35	0.35	0.37	0.37	0.24	0.23	0.27	0.27
Municipalities	301	301	373	373	301	301	373	373
Mun. × Time	2107	2107	2606	2606	2106	2106	2605	2600
Panel B: The Russian - Prussian border								
Partitions effect	−.149	−.105	−.099	−.148	−.235	-.011	0.02	−.06
(Russia = 1)	(.188)	(.309)	(.13)	(.127)	(.213)	(.313)	(.145)	(.14)
Interaction term	.00008	0.0003	−.00002	0.0001	.00006	−.003	−.00009	−.0001
	(.0002)	(.004)	(.00003)	(.0002)	(.0001)	(.004)	(.00004)*	(.0001)
Mean of outcome	-.22	−.22	−.15	−.15	−.04	−.04	−.03	−.03
R^2	.1	0.11	0.14	0.14	0.16	0.17	0.14	0.14
Municipalities	206	206	302	302	206	206	302	302
Mun. × Time	1442	1442	2114	2114	1442	1442	2114	2114
Interaction var.	Pop.	Dens.	Pop.	Dens.	Pop.	Dens.	Pop.	Dens.
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Modern controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Rural	Rural	All	All	Rural	Rural	All	All

Notes: Robust and clustered standard errors (at municipality level) 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-stakes exam score; Columns 5 to 8 - the mathematics and science 9th grade high-stakes 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. In addition the regressions include an interaction term between the partition dummy and population number (in 10 thousand) (Columns 1, 3, 5 and 7) or population density (Columns 2, 4, 6 and 8), a quadratic polynomial in latitude and longitude and a set of geographic and modern covariates. All the regressions use 50 km bandwidth.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jce.2018.10.007

References

Abramitzky, R., Boustan, L.P., Eriksson, K., 2014. A nation of immigrants: assimilation and economic outcomes in the age of mass migration. J. Political Econ. 122 (3), 467–506.
Acemoglu, D., Johnson, S., Robinson, J.A., 2001. The colonial origins of comparative development: an empirical investigation. Am. Econ. Rev. 91 (5), 1369–1401.
Acemoglu, D., Robinson, J.A., 2008. Persistence of power, elites and institutions. Am. Econ. Rev. 98 (1), 267–293.
Akerlof, G., Kranton, R., 2002. Identity and schooling: some lessons for the economics of education. J. Econ. Lit. 40 (4), 1167–1201.
Akerlof, G., Kranton, R., 2010. Identity Economics: How Our identities Shape our Work, Wages, and Well-Being. Princeton University Press, Princeton, NJ.
Alesina, A., La Ferrara, E., 2005. Ethnic diversity and economic performance. J. Econ. Lit. 43 (3), 762–800.
Angrist, J.D., Lavy, V., 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. Q. J. Econ. 114 (2), 533–575.
Angrist, J.D., Pischke, J.-S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, NJ.
Backhaus, A., 2018. Fading legacies: human capital in the aftermath of the partitions of poland. mimeo.
Banerjee, A., Iyer, L., 2005. History, institutions, and economic performance: the legacy of colonial land tenure systems in india. Am. Econ. Rev. 95 (4), 1190–1213.
Basten, C., Betz, F., 2013. Beyond work ethic: religion, individual, and political preferences. Amer. Econ. J. 5 (3), 67–91.
Becker, S.O., Boeckh, K., Hainz, C., Woessmann, L., 2016. The empire is dead, long live the empire! long-run persistence of trust and corruption in the bureaucracy. Econ. J. 126 (590), 40–74.
Becker, S.O., Cinnirella, F., Hornung, E., Woessmann, L., 2014. Ipehd–the ifo prussian economic history database. Hist. Method. 47 (2), 57–66.
Becker, S.O., Woessmann, L., 2009. Was weber wrong? a human capital theory of protestant economic history. Q. J. Econ. 124 (2), 531–596.
Bisin, A., Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. J. Econ. Theory 97 (2), 298–319.
Bowen, J., 1962. Soviet Education: Anton Makarenko and the Years of Experiment. University of Wisconsin Press, Madison, WI.
Burke, P., 1979. The new cambridge modern history: Volume 13, companion volume. 13 Cambridge University Press, Cambridge.
Cantoni, D., Yuchtman, N., 2013. The political economy of educational content and development: lessons from history. J. Dev. Econ. 104, 233–244.
Caselli, F., Coleman, W.J., 2013. On the theory of ethnic conflict. J. Eur. Econ. Assoc. 11 (s1), 161–192.
Cassar, A., d'Adda, G., Grosjean, P., 2014. Institutional quality, culture, and norms of cooperation: evidence from behavioral field experiments. J. Law Econ. 57 (3), 821–863.
Chubarov, A., 2000. The fragile empire: A history of imperial russia. Blackwell Publisher, Hoboken NJ.
Chuminski, J., 2008. Gospodarcze i spoleczne skutki zaborow polski. GAJT, Wroclaw.
Cinnirella, F., Schueler, R.M., 2016. The cost of decentralization: linguistic polarization and the provision of education. CESifo Working Paper Series No. 5894.
Cohen, G.B., 1996. Education and middle-class society in imperial austria, 1848–1918. Purdue University Press, West Lafayette, IN.
Cvrcek, T., Zajicek, M., 2013. School, what is it good for? useful human capital and the history of public education in central europe. NBER Working Paper No. 19690.
Czapiński, J., Panek, T., 2013. Social diagnosis 2011 - 2013. http://www.diagnoza.com [October 2015];.
Davies, N., 2005. God'S playground a history of poland: Volume i: The origins to 1795. 1 Oxford University Press, Oxford.
Davies, N., 2005. God'S playground a history of poland: Volume II: 1795 to the present. 2 Oxford University Press, Oxford.
Dell, M., 2010. The persistent effects of peru's mining mita. Econometrica 78 (6), 1863–1903.
Dell, M., Lane, N., Querubin, P., 2018. The historical state, local collective action, and economic development in vietnam. Econometrica (forthcoming).
Denny, K., Doyle, O., 2008. Political interest, cognitive ability and personality: determinants of voter turnout in britain. Br. J. Political Sci. 38 (2), 291–310.
Dippel, C., 2014. Forced coexistence and economic development: evidence from native american reservations. Econometrica 82 (6), 2131–2165.
Dobrowolski, T.N., 1886. Biskupstwo tarnowskie. rys historyczny. Biblioteka Warszawska 3, 365–373.
Dolata, R., Hawrot, A., Humenny, G., Jasińska, A., Koniewski, M., Majkut, P., Żółtak, T., 2013. Trafność metody edukacyjnej wartości dodanej dla gimnazjów. Instytut Badań Edukacyjnych.
Dolata, R., Hawrot, A., Humenny, G., Jasińska-Maciążek, A., Koniewski, M., Majkut, P., Blukacz, M., Grygiel, P., Otręba-Szklarczyk, A., 2015. (Ko) warianty efektywności nauczania. wyniki badania w klasach IV–VI. Instytut Badań Edukacyjnych.
Feir, D., 2016. The intergenerational effect of forcible assimilation policy on education. Int. Indigenous Policy J. 34 (2), 185–196.
Galor, O., Moav, O., 2006. Das human-kapital: a theory of the demise of the class structure. Rev. Econ. Stud. 73 (1), 85–117.
Gelman, A., Imbens, G., 2018. Why high-order polynomials should not be used in regression discontinuity designs. J. Bus. Econ. Stat. 0 (0), 1–10.
Gillis, J.R., 1971. The prussian bureaucracy in crisis, 1840–1860: origins of an administrative ethos. Stanford University Press, Palo Alto, CA.
Goldin, C.D., Katz, L.F., 2009. The race between education and technology. Harvard University Press, Cambridge, MA.
Grabowska, M., 2004. Podział postkomunistyczny: Społeczne podstawy polityki w polsce po 1989 roku. Wydawnictwo Naukowe Scholar.
Graham, J., Haidt, J., Nosek, B.A., 2009. Liberals and conservatives rely on different sets of moral foundations. J. Pers. Soc. Psychol. 96 (5), 1029.
Greenstein, F.I., 1969. Children and politics. Yale University Press, New Haven, CT.
Grosfeld, I., Rodnyansky, A., Zhuravskaya, E., 2013. Persistent antimarket culture: a legacy of the pale of settlement after the holocaust. Amer. Econ. J. 5 (3), 189–226.
Grosfeld, I., Zhuravskaya, E., 2015. Cultural vs. economic legacies of empires: evidence from the partition of poland. J. Comp. Econ. 43 (1), 55–75.
GUS, 1938. Drugi powszechny spis ludności z dnia 9 grudnia 1931 roku. mieszkania i gospodarstwa domowe. ludność. stosunki zawodowe.
GUS, 2003. Historia polski w liczbach - Panstwo i spoleczenstwo. GUS Warszawa.
GUS, 2014. Historia polski w liczbach - Polska w europie. GUS Warszawa.
Herbst, M., 2004. Zroznicowanie jakosci kapitalu ludzkiego w polsce. od czego zaleza wyniki edukacyjne? Studia Regionalne i Lokalne 17 (3), 93–104.
Herbst, M., Herczynski, J., Levitas, A., 2009. Finansowanie oswiaty w polsce: Diagnoza, dylematy, mozliwosci. Wydawnictwo Naukowe Scholar.
Herbst, M., Rivkin, S., 2012. Divergent historical experiences and inequality in academic achievement: the case of poland. J. Socio. Econ. 42, 1–12.
Herczynski, J., Sobotka, A., 2013. Siec szkol podstawowych i gimnazjow w Polsce 2007–2012. Technical Report. Instytut Badań Edukacyjnych.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., et al., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25 (15), 1965–1978.
Hryniewicz, J.T., 2003. Polityczne i gospodarcze nastepstwa zaboru pruskiego. Studia Regionalne i Lokalne 13 (3), 61–78.
Huillery, E., 2009. History matters: the long-term impact of colonial public investments in french west africa. Amer. Econ. J. 1 (2), 176–215.
Ilski, Z., 2008. Oboz narodowodemokratyczny wobec zasad i formuly wyborczej w okresie poprzedzajacym odbudowe niepodleglej polski. Slaski Kwartalnik Historyczny Sobotka 63 (2).
Imbens, G.W., Lemieux, T., 2008. Regression discontinuity designs: a guide to practice. J. Econom. 142 (2), 615–635.
Iyer, L., 2010. Direct versus indirect colonial rule in india: long-term consequences. Rev. Econ. Stat. 92 (4), 693–713.
Jha, S., 2013. Trade, institutions, and ethnic tolerance: evidence from south asia. Amer. Political Sci. Rev. 107 (04), 806–832.
Karaja, E., 2013. The Rule of Karlowitz. Fiscal Change and Institutional Persistence. Technical Report. University of Berkeley, BEHL Working Paper No. 2014-04.
Kashin, K., Ziblatt, D., 2012. Administrative district boundaries, germany, 1890s. Cambridge, Mass.: Harvard Geospatial Library. Retrieved from http://nrs.harvard.edu/urn-3:HUL.gisdata:013361293.
Kassof, G., 2004. Education. Encycl. Russian Hist. 500.
Keele, L.J., Titiunik, R., 2014. Geographic boundaries as regression discontinuities. Political Anal. 23 (1), 127–155.

Kowalski, M., 2000. The electoral geography of poland: spatial differences in electoral behaviour, 1989–1998. Polish Academy of Sciences, Warszawa: Institute of Geography and Spatial Organization, Geopolitical Studies 7.
La Porta, R., Lopez-de Silanes, F., Shleifer, A., Vishny, R., 1999. The quality of government. J. Law, Econ. Org. 15 (1), 222–279.
Labuda, G., 1971. Polska granica zachodnia: Tysiac lat dziejow politycznych. Wydawnictwo Poznanskie.
Lamberti, M., 1989. State, society, and the elementary school in imperial germany. Oxford University Press, New York, NY.
Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48 (2), 281–355.
Lee, J., Grigg, W., Dion, G., 2007. 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.
Meissner, A., Majorek, C., 2000. Pedagogika nowego wychowania w polsce u schylku XIX iw pierwszej polowie XX wieku: Podstawowe przejawy i wspolczesne odniesienia. Wydawictwo Wyzszej Szkoly Pedagogicznej, Rzeszów.
Michalopoulos, S., Papaioannou, E., 2013. Pre-colonial ethnic institutions and contemporary african development. Econometrica 81 (1), 113–152.
i Miquel, G.P., 2007. The control of politicians in divided societies: the politics of fear. Rev. Econ. Stud. 74 (4), 1259–1274.
Moon, D., 1992. The determinants of turnout in presidential elections: an integrative model accounting for information. Political Behav. 14 (2), 123–140.
MPIDR [Max Planck Institute for Demographic Research], CGG [Chair for Geodesy and Geoinformatics, University of Rostock], 2012. MPIDR population history GIS - Rostock.
Patacchini, E., Zenou, Y., 2011. Neighborhood effects and parental involvement in the intergenerational transmission of education. J. Reg. Sci. 51 (5), 987–1013.
Putnam, R.D., Leonardi, R., Nanetti, R.Y., 1994. Making democracy work: Civic traditions in modern italy. Princeton University Press, Princeton, NJ.
Sakalli, S.O., 2018. Secularization and religious backlash: evidence from turkey. mimeo.
Snyder, T., 2006. "Ukrainians and poles" in: The cambridge history of russia: Volume 2, imperial russia, 1689–1917. 2 Cambridge University Press.
Spolaore, E., Wacziarg, R., 2016. Fertility and modernity. UCLA CCPR Population Working Papers.
Steele, C.M., Aronson, J., 1995. Stereotype threat and the intellectual test performance of african americans. J. Pers. Soc. Psychol. 69 (5), 797–811.
Sulimierski, F., Chlebowski, B., Wlalewski, W., 1895. Słownik geograficzny królestwa polskiego i innych krajów słowiańskich. 14 Filipa Sulimierskiego i Władysława Walewskiego.
Trzeciakowski, L., 2003. Poslowie polscy w berlinie: 1848–1928. Wydawnictwa Sejmowe, Warsaw.
Urquiola, M., 2016. Competition among Schools: Traditional Public and Private Schools. Handbook of the Economics of Education. 5. Elsevier, pp. 209–237.
Wandycz, P.S., 1974. The lands of partitioned poland, 1795–1918. 7 University of Washington Press.
Whittaker, C.H., 1984. The origins of modern russian education: An intellectual biography of count sergei uvarov, 1786–1855. Northern Illinois University Press, DeKalb, IL.
Wojciak, J., 1981. Walka polityczna w wyborach do parlamentu rzeczy i sejmu pruskiego w poznanskiem w latach 1898–1914. Panstwowe Wydawnictwo Naukowe, Warsaw.
Wolf, N., 2007. Endowments vs. market potential: what explains the relocation of industry after the polish reunification in 1918? Explor. Econ. Hist. 44 (1), 22–42.
Wysokinska, A., 2011. Importance of History for Economic Development Lessons from Natural Experiment of History. European University Institute.
Zarycki, T., 2007. History and regional development. a controversy over the 'right'interpretation of the role of history in the development of the polish regions. Geoforum 38 (3), 485–493.
Zubrzycki, J., 1953. Emigration from poland in the nineteenth and twentieth centuries. Popul. Stud. (Camb.) 6 (3), 248–272.