10.14754/CEU.2016.03

ESSAYS ON THE ECONOMICS OF EDUCATION

by

Paweł Bukowski

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at Central European University

Advisor: Gábor Kézdi

Associate Advisor: John S. Earle

Budapest, Hungary

10.14754/CEU.2016.03

CENTRAL EUROPEAN UNIVERSITY
DEPARTMENT OF ECONOMICS

The undersigned hereby certify that they have read and recommend to the Department of Economics for acceptance a thesis entitled **"Essays on the Economics of Education"** by **Pawel Bukowski**
Dated: June 20, 2016

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Chair of the Thesis Committee:

László Csaba

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Advisor:

John Earle

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Advisor:

Gábor Kézdi

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Internal Examiner:

Sergey Lychagin

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
External Examiner:

Sascha O. Becker

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
Internal Member:

Miklós Koren

10.14754/CEU.2016.03

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

External Member:

Júlia Varga

I certify that I have read this dissertation and in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

External Member:

Dániel Horn

10.14754/CEU.2016.03

CENTRAL EUROPEAN UNIVERSITY
DEPARTMENT OF ECONOMICS

Author: Paweł Bukowski
Title: Essays on the Economics of Education
Degree: PhD
Date: 20 June 2016

Hereby I testify that this thesis contains no material accepted for any other degree in any other institution and that it contains no material previously written and/or published by another person except where appropriate acknowledge is made.

Signature of the author: ....
Paweł Bukowski

10.14754/CEU.2016.03

DISCLOSURE OF CO-AUTHORS CONTRIBUTION

Title of paper: The Threat of Competition and Public School Performance: Evidence from Poland

Co-author: Martyna Kobus

The nature of cooperation and the roles of the individual co-authors and approximate share of each co-author in the joint work: The paper was developed in cooperation with Martyna. My contribution was more pronounced in working out the methodology, data collection and running the regressions, while Martyna's contribution was more pronounced in writing the text of the paper.

10.14754/CEU.2016.03

Abstract

This thesis consists of one co-authored and two single-authored chapters, which investigate the modern and historical determinants of student performance and educational inequalities.

The first chapter (and the Job Market Paper) examines the effect on current student performance of the 19th century Partitions of Poland among Austria, Prussia and Russia. Despite the modern similarities of the three regions, using a regression discontinuity design I show that student test scores are 0.61 standard deviation higher on the Austrian side of the former Austrian-Russian border. On the other hand, I do not find evidence for differences on the Prussian-Russian border. Using a theoretical model and indirect evidence I argue that the Partitions have persisted through their impact on social norms toward local schools. Nevertheless, the persistent effect of Austria is puzzling given the historical similarities of the Austrian and Prussian educational systems. I argue that the differential legacy of Austria and Prussia originates from the Austrian Empire's policy to promote Polish identity in schools and the Prussian Empire's efforts to Germanize the Poles through education.

The second chapter is co-authored with Martyna Kobus. It studies the effect of threat of school competition on performance of public schools. We provide the evidence for the negative effect of the competition on students' test scores in public schools. We use the introduction of the amendment facilitating the creation of autonomous schools in Poland in 2009 as a breakthrough date in DiD estimation. The specifics of the Polish reform provide for a credible proxy for the *threat* of competition, so we can take into account that the size of competition is endogenous to the market characteristics. For the total sample we find no effect, however, for more competitive urban educational markets, we report a drop in test scores in public schools following the introduction of the amendment. This negative effect is robust to the existence of some competition prior to the amendment and to the size of public schools. It does not result from the violation of the common trend assumption either. We focus on the short run in which there is only a limited set of actions available to schools' principals. We exclude student sorting as a potential channel.

The third chapter estimates the effect of school competition on sorting within a school (across classes). The identification strategy is based on a two-stage design of the Polish Comprehensive Education, which allows to isolate an exogenous change in student mobility. In addition, I use a novel measure of student socio-economic characteristics - Raven's Progressive Matrix test score. The results show that school competition leads to a higher sorting of students within a school and between schools. I investigate two explanation of the effect on sorting within a school: the demand for peer quality and the demand for teachers. The data point to the importance of the former mechanism, i.e. the demand for high quality peers that motivates school principals to create high tracks within a school.

10.14754/CEU.2016.03

Chapter 1

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

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. In the case of education, historical investments in public goods and property rights institutions have been shown to affect current educational attainment, provision of schools and literacy levels. However, we know less about the mechanisms underlying these long-run consequences of institutions and if and how they depend on social context. In the first chapter, I analyze the Partitions of Poland (1815-1918) among Austria, Prussia, and Russia as a laboratory to investigate how history matters for student performance. The existing evidence suggest that the former borders between the Empires were not drawn to reflect any pre-existing socio-economic, historical, geographic or ethnic divisions. Today, the the three regions are within Poland, are ethnically homogeneous and have the same modern educational and legal systems.

Using a two-dimensional regression discontinuity design I compare test-measured performance of students in municipalities at the two sides of the former border between Austria and Russia. I show that the municipality-average student test scores on the Austrian side are 0.61 standard deviation higher. On the other hand, I do not find evidence for differences on the Prussian-Russian border.

There are many potential channels through which the Partitions has affected the current student performance. I highlight that people living in the former Austrian Empire have inherited positive social norms toward local schools, which lead to a higher schooling effort and thus increase the performance of students. I provide three pieces of empirical evidence to support this channel. At the same time, I show that other channels, in particular skill-biased migrations, labor market differences, school quality are unlikely to explain my results.

Why social norms differ in the Austrian and Prussian partitions is puzzling given given that the former was not economically superior over the later and both Empires had almost identical educational systems and similar provision of public education. I argue that the differential legacy of Austria and Prussia originates from the different interaction between educational institutions and Polish identity. While the Prussian state used these institutions mainly to Germanize Poles (e.g. through the German language of instruction), the Austrian state used them to support Polish identity (e.g. through the Polish language of instruction). Because of the historical attitude of the Polish population toward the educational systems, positive social norms toward education may have been more likely to emerge in the Austrian partition. These could be then transmitted through generations and still affect student and parental effort. I provide a suggestive evidence for this hypothesis using the historical data on the 19th century educational outcomes

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

Chapter 2

”The threat of competition and public school performance: evidence from Poland”

with Martyna Kobus

School competition, one of the most important forces driving effectiveness of schools, depends on school funding, school choice or the structure of education. However, schools are not only affected by the actual competition, they might also anticipate a threat of competition and try to prevent it. Similarly, as monopolistic firms, they might want to block entry of new schools in order to secure their profits. But this might not be beneficial for student performance.

In the second chapter we focus on the impact of the threat of competition from autonomous schools on public school performance in Poland. As an identification strategy we use the amendment to the education act introduced in March 2009 which facilitated the creation of autonomous schools but only for schools that have 71 and less students. Therefore, public schools located in areas where there is a higher percentage of students who attend schools with less than 71 students are more exposed to competition. We show that this is indeed strongly related to actual creations of autonomous schools. Using year 2009 as a breakthrough date in the Difference-in-Difference estimations we find that the higher competition caused by the mentioned reform has significant negative impact on the performance of urban public schools. Urban areas are more competitive educational markets than rural areas, with dense school network and better parental background. The effect is similar for public schools that are larger (more than 300 students) and becomes stronger for urban schools that already have an autonomous school in their neighborhood and may thus be more aware of the consequences of the reform. We argue that in our case changes in student composition between schools are unlikely, but also we analyze changes in test scores at the municipality level, which cancels out the direct effects of sorting between schools. In cities (above 20000 inhabitants) test scores drop at the aggregate community level.

We focus on the short run consequences of the reform. In the short run, school administrators are restricted in their options and school’s reputation is not only the function of productivity and student ability, but also on activities that are visible to parents and attract but are not necessarily related to productivity. In Poland, teachers enjoy high level of employment security and they cannot be laid off easily, in particular, not in the short run. Therefore, what remains available to school principals, is either efficiency changes i.e. incentivizing teachers to work harder, or boosting their school’s prestige. We find no evidence for increasing teachers’ salaries or investment in infrastructure in the expenditure dataset. We only find some anecdotal evidence that principals may resort to deterring the entry of community schools by marketing their schools to parents.

Chapter 3

“School Competition and Sorting of Students Within a School”

Existing literature shows that school competition might lead to sorting of students between schools. However, we know less about the effect of school competition on sorting within a school (across classes). This is surprising given the importance of class assignment and that student sorting is not neutral for the performance of students and might violate educational equality of opportunity.

This chapter is a first study to estimate the effect of school competition on sorting within a school (across classes) and between schools. In order to isolate an exogenous change in student mobility, which increases school competition, I exploit a two-stage design of the Polish Comprehensive education. For measuring sorting I use the fraction of the variance of Raven’s Progressive Matrix test score explained by school or class levels. Raven’s score is a measure of general intelligence, which is determined by student genetic abilities and socio-economic background. It is fixed since early childhood, which ensures that the only source of class/school homogeneity is sorting of students. The results show that school competition leads to more homogeneous classes and schools.

Next, I focus on the potential mechanisms explaining the effect of school competition on sorting within a school. One explanation is that high track might be used to attract high-skill or high-income students (the demand for peer quality). A complementary is that it might be also used to attract high-skilled teachers (the demand for teachers). Using data on school characteristics I empirically test these two channels. The results point out to the importance of the demand for peer quality channel.

10.14754/CEU.2016.03

Acknowledgements

First and foremost, I am indebted to my advisor Gábor Kézdi and associate advisor John S. Earle for all the valuable conversations, perpetual support and inspiring criticism. I am especially glad for having worked with Martyna Kobus, Filip Novokmet, Gregory Clark, Attila Gáspár and Rita Pető with whom I experienced the miracle of teamwork.

I am grateful to my examiners, Sascha O. Becker and Sergey Lychagin, for their useful comments and encouragement.

I would like to express my sincere thanks to my professors, fellow students and staff at the Central European University, who created an exciting and stimulus working environment. I am especially thankful to Miklós Farkas, who was not only a critical reader of this dissertation but also a great friend of mine. My colleagues from the PhD room, Márta Bisztray, Gergely Hajdú, Balint Menyhert, Balázs Reizer, Gábor Revesz, István Szabó and Peter Zsohár provided me with excellent suggestions and everyday joy.

I gratefully acknowledge financial support from the CEU foundation throughout my study and for my visit at the University of California at Berkeley, where I received valuable comments from David Card, Bernhard Enzi, Hedvig Horvath, Attila Lindner and Hilda Ralsmark.

Last but not least, I am especially thankful to my family for their support, patience and love. I owe Anna Mączyńska for her patience and encouragement throughout these years. I have received a great and constant support from my friends and I am especially indebted to Adam Araszkiewicz, Silvija Aurylaitė, Petra Baji, Olga Cojocaru, Ania Fundowicz, Ola Gawryś, Ania Godziszewska, Emöke Győrfi, Maciej Kowalczyk, Miłosz Waligórski, Jonathon Winkler, Karolina Wnuk, Justyna Wrzosek and Ola Zagrodzka.

10.14754/CEU.2016.03

Table of Contents

10.14754/CEU.2016.03

Chapter 1
How History Matters for Student Performance. Lessons from the Partitions of Poland *

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

*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 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. All errors are mine. The previous version of this paper was publicized through the Berkeley Economic History Lab Working Papers series.

10.14754/CEU.2016.03

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 institutions have been shown to affect current educational attainment, 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 educational outcomes (Dell, 2010). In this paper, I show that two historical parts of Poland, which had similar past educational system and provision of public education, had - relative to a control region - very different long run effects on current student performance. I show evidence highlighting the role of social norms toward local schools as a key channel of persistence (Akerlof and Kranton, 2010; Sakalli, 2014). I also argue that the interaction between national identity and institutions created different social norms toward local schools in the two historical parts, generating the difference in student performance today.

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

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

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

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

Why social norms differ in the Austrian and Prussian partitions is puzzling given given that the former was not economically superior over the later and both Empires had almost identical educational systems and similar provision of public education.[2] I

[2]Consistently, Michalopoulos and Papaioannou (2013) and Grosfeld and Zhuravskaya (2015) show that

argue that the differential legacy of Austria and Prussia originates from the different interaction between educational institutions and Polish identity. While the Prussian state used these institutions mainly to Germanize Poles (e.g. through the German language of instruction), the Austrian state used them to support Polish identity (e.g. through the Polish language of instruction) (Cohen, 1996; Lamberti, 1989).[3] Because of the historical attitude of the Polish population toward the educational systems, positive social norms toward education may have been more likely to emerge in the Austrian partition. These could be then transmitted through generations and still affect student and parental effort. Consistent with this hypothesis, Steele and Aronson (1995) and Akerlof and Kranton (2010) provide theoretical and empirical evidence that identity is associated with social norms affecting an individual's schooling choices, school-student relationships and student achievements.

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

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

not all institutions influence social norms and matter in the long run.

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

interaction between institutions and identity.

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

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

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

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

10.14754/CEU.2016.03

2 Historical Overview

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

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

The new border between the partitions was established during the Congress of Vienna in 1815 after which they remained generally unchanged until the end of World War I. In this section, by the Austrian Partition I mean the areas which correspond to the historical region of Galicia, that is, the southern parts of modern Poland, along with 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. Therefore, cities such as Wrocław (*ger.* Breslau) or Gdańsk (*ger.* Danzig) 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 the modern Belarus and Ukraine. In the empirical part of the paper, the definition of each Partition is limited only to the territories within the modern Poland.

During the first decades after the Congress, the Russian and Prussian administrations were not systematically oppressive toward the Poles. The Congress Kingdom and the

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

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

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

The best economic situation was in the Prussian zone. The authorities carried out

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

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

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

many reforms there. The most important of these was the manumission, that is the peasants could become owners of the land after repaying the nobility. Rising demand for agricultural products induced changes in agricultural technology (crop rotation), fertilizers were applied, and the wealthier farmers were buying machinery. As a result, the agriculture, rather than industry, was the main drive of the economic progress in the Grand Duchy of Poznań. Economies of the other partitions were different. In the Russian zone it was industry that developed the most. The clusters of textile industry were created in Łódź and Białystok. Warsaw became a modern city with its sewers, streets, gas lighting, and power plant switchboard. Economic progress, however, did not improve the well-being of workers who had to work long hours (14 hours) for low wages and in unsafe conditions. The delayed manumission reforms, which were introduced only during the second half of the 19th century, contributed to the relative backwardness of the agriculture in the Congress Kingdom. Nevertheless, the worst economic situation was in the Austrian part. Before the end of the 19th century Galicia had not been industrialized and the agriculture was under invested and parceled. Consequently, people had experienced one of the worst poverty rates in all of the Habsburg Empire, and at the beginning of the 20th century over two million Galicians emigrated abroad to escape the bad economic conditions.

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

2.1 The Prussian Education System

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

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

Nevertheless, the policy turned out to be very important in the Polish context. The local Catholic Church[10] helped to cultivate the Polish national identity more than any other secular movement. Consequently, *kulturkampf* was done more consistently on the Polish lands than anywhere else (Lamberti, 1989). In addition to this, from 1870 the Prussian state executed repressions on a much larger scale than it had done before, in particular, it banned the use of the Polish language in administration and education, forbidden cultivation of the Polish traditions, discriminated Polish workers on the labor market and deported Poles and Jews, who did not have the Prussian citizenship.

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

[9]Understood as the German nation.

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

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

- retained the policy. "This policy bred germanophobia and a repugnance for the school in Polish families" (Lamberti, 1989, p.109). But the language of instruction was not the only reason why Polish parents opposed the educational system.

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

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

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

The German language of instruction, inequalities, the role of teachers and the interconfessional education motivated the hostility toward the education system among the

[12]Still it was much better than in Russia or Austria, see for example illiteracy rates in Figure 2.
[13]Interconfessional schools (also called mixed) gather students from different religious groups.

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

2.2 The Austrian Education System

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

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

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

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

The class instruction in Polish, broadening the access to education and poverty caused

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

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

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

A remaining question, for which answer is beyond the scope of this paper, is why the Prussian and Austrian Empires had different ethnic policy since 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, initially the Habsburg Empire was less tolerant, but the Revolutions of 1848 contributed to the reorientation of their ethnic policy. 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

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

The other problem was the chaotic organization of the school system. There was no obligatory schooling and the educational policy was inconsistent, as the Ministry of Education did not control the network of schools.[17] Lacking the central design and organization, the system was characterized by class-based duality, with separate curricula for students from upper and lower stratum. Consequently, the illiteracy levels were very

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

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

high: in 1917, only 70% and 30% of urban and rural population respectively could read and write.[18]

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

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

2.4 Summary

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

In terms of educational systems, the Austrian and Prussian institutions were very similar as the former was copied from the latter. The Austrian and Prussian system was financed from local taxes, had compulsory elementary and optional secondary education

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

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

10.14754/CEU.2016.03

and shared similar curricula and pedagogical methods. However, while the Prussian state used these institutions mainly to Germanize Poles (e.g. through the German language of instruction), the Austrian state used them to support Polish identity (e.g. through the Polish language of instruction). Consequently, the Poles under Prussian rule opposed the educational system and were hostile toward the school personnel (especially teachers). Remarkably, massive school strikes were organized by Polish parents, the largest one took place in 1901 when 70 thousands Polish students refused to go to school. The Russian educational system, in turn, practically did not exist in the 19th century.

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

Today, the Polish comprehensive and compulsory education system consists of 6 years of elementary school, which is then followed by 3 years of *gimnazjum*. The admission to the comprehensive schools is based on catchment areas, which means that every student living within this area has a right to attend a given public school. However, parents may request an alternative school, but its principal has a right to reject the application. Next 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 the admission is not granted.

10.14754/CEU.2016.03

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 stake general 6th grade exam (taking place after elementary school) and a high stake mathematics-science 9th grade exam (after lower secondary school). While the former serves mainly an informational purpose, the later matters in the high school admission process and thus motivates students (and their parents) to obtain the best score. The tests are standardized for all of Poland and are corrected outside school by randomly chosen professional test checkers. Students do not know their identity and vice versa.

In addition to this, a set of socio-economic control variables are available at the municipality level from the Central Statistical Office of Poland and the System of Educational Information. Geographical and climate data come from the *WorldClim.org* project (Hijmans, Cameron, Parra, Jones, Jarvis et al., 2005). For the full description of the available variables see Table A1.

Descriptive statistics for *rural* municipalities located at most 50km from the borders are presented in Table 3. These variables are for the present period, and, as such, are endogenous with respect to the Partitions of Poland.[20] The border areas of the former Russian Partition seem to have the worst socio-economic situation, as the rate of unemployment is the highest, the expenditures are the lowest and the migration balance is negative. Importantly, the former Russian municipalities on the Austrian-Russian and Prussian-Russian borders are similar (columns (3) and (8)). The municipalities which were under the Prussian rule are characterized by high share of employment in agriculture [21], high share of people aged 0-18, positive migration balance and low level of unemployment. The situation in the former Austrian zone is similar to the former Prussian, except a low importance of agriculture and high population number and density.

The rural borderlands of the former Austrian partition has the best educational outcomes (except the number of additional classes and the level of scholarization), even

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

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

though these lands are not necessary better in the case of the other socio-economic characteristics. Importantly, it has also higher educational spending per capita, but this difference disappears when the general spatial trends are accounted for (see Table 6). The former Prussian and Russian borderlands have similar level of achievements, but the former have the largest classes and highest number of additional lessons. The high performance of students from the former Austrian partition is also confirmed by Figures 3 and 4, which show the spatial distribution of the 2011 6th and 9th grade exams for the whole sample of municipalities. It can be clearly seen that the territory of the former Austrian Empire is a cluster of high-performing municipalities.

3.2 Empirical Strategy

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

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

where i indexes municipality and t indexes year. $f(location_i)$ is a polynomial of latitude and longitude, the dummy D takes value 1 for the former Russian areas and value 0 for either the Austrian or Prussian, G_i are time-invariant geographical controls (altitude, precipitation and temperature), and ϵ_{it} denotes idiosyncratic shocks. The two outcome variables are the standardized (Z-score) 6th grade exam score and the standardized mathematics and science 9th grade exam scores.[23] They are available from

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

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

2005 to 2011. The sample consists of municipalities, which are located within a given distance to the borders (the bandwidth).[24] I pool the data and estimate the model using the Random Effect estimator 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 the OLS estimates for each year separately.

The regression discontinuity framework requires a proper specification of the polynomial $f(location_i)$ and the bandwidth. There are no theoretical arguments for any specific order, therefore I report results for linear, quadratic, cubic and quartile polynomials. Table A4 reports the Akaike Information Criteria, which can be used for the model selection (Lee and Lemieux, 2010). The quadratic and quartile polynomials are favoured, but the quadratic polynomial has more degrees of freedom and I use it in a baseline specification. Nevertheless, 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 choose 50km bandwidth and in the robustness section I also report results for municipalities located at most 75km and 100km.

3.3 The Borders under Investigation

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

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

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

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

ing the borders under investigation, the dashed line marking the excluded parts. Please note that by excluding Silesia I cannot directly compare the borderlands between the Austrian and Prussian partitions. The bottom part of Figure 1 presents the total area under investigation, namely the rural and urban municipalities located at most 50km from the borders of interest.

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

The Partitions borders under investigation were mostly set along rivers. Between the Prussian and Russian Empires it was drawn along the Drwęca and Prosna rivers (which are small waterways), whereas the half of the Austrian - Russian border was drawn along the Vistula river. Becker et al. (2014b) show that there are no significant differences between these regions in terms of geography and pre-Partitions historical characteristics. Similarly, Grosfeld and Zhuravskaya (2015), using the one-dimensional non-parametric regression discontinuity design, report only a small difference in altitude on the Austrian-Russian border. Yet, my own estimations in Table 4 Panel A show that there are also significant differences in temperature and precipitation on the Austrian-Russian border when the two-dimensional specification is used. The "jump" in altitude on the Russian side of the Austrian-Russian border is around 80 meters, precipitation is higher by 30mm and temperature drops by around 0.45 C°. The magnitudes are not large and they arise most likely because of the riverbed of the Vistula. On the Prussian-Russian border the two-dimensional specification (Table 4 Panel C) also reports significant differences, but with smaller magnitudes.[27] Still, in order to control for the potential confounding effect, I control for the geographic and climate characteristics in the baseline regressions (the estimates are generally insensitive to their inclusion). 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 the educational differences between the borderlands or induce other dissimilarities in culture or institutions.

When the border exogeneity assumption hold, estimation of the discontinuous change

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

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

3.4 Results

Figures 5 and 6 present relationship between the average student performance and distance to the Austrian-Russian and Prussian-Russian borders respectively. A drop in the 6th and 9th grade exam score can be seen clearly at the border between Austria and Russia (positive distance) indicating a strong and positive effect of the former Austrian Empire. Contrary to this, no visible effect can be seen on the Prussian-Russian border.

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

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

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

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

controls are included. Contrary to the Austrian-Russian border, these results show that students from the former Prussian zone do not perform better than those from the former Russian territories. In fact, the estimated absolute effects of the Austrian Empire on the Russian-Austrian border are significantly larger (at the 0.1% level) from the effects of the Prussian Empire on the Russian-Prussian border (the comparison is not reported).

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

3.5 Robustness

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

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

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

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

border, the magnitudes increase in absolute terms but they are still insignificant (with an exception of column (1)).

The baseline results might be sensitive to the specification choices. Table A5 reports estimates of the Partitions effects for different polynomials in latitude and longitude, along with different bandwidth choices. All regressions include the geographic 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. Contrary to this pattern, the estimates of the Partitions effect on the Prussian-Russian border vary a lot across specifications. Importantly, the sign changes once the bandwidth is increased to 75km and 100km - which indicate that students living in the Russian zone are, in fact, perform better than those from the Prussian one. These contradict the findings from Table 5. Nevertheless, in most of the cases the coefficients are not significant.

The same set of specification choices is examined with the total sample (Table 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. Results for the Prussian Empire are similar as previously.

In order to see whether the results are driven by a particular year, I estimate the baseline Equation (1) by OLS for each year separately and the rural sample. Figure 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 stake exam and shifted toward zero in the case of the 9th grade high stake exam. By contrast, the effect of the Prussian Empire is null across years and types of the exam.

Next, I check whether the results are sensitive to an inclusion of a set of time-variant modern controls. The variables reflect general educational differences (local government expenditures per capita on education, kindergarten and secondary school attendance) and the socio-economic conditions (local government total expenditures per capita, unemployment ratio, population level, population density and migration balance). Table A1 provides the exact definition of the variables. Importantly, note that these covariates are endogenous, that is, they are also affected by the Partitions of Poland (see Section 4.3 and Table 6). As such, the augmented regression "switches off" some potential channels of influence of the Partitions of Poland on outcomes and leads to "bad control bias"

(Angrist and Pischke, 2008).[31] Indeed, as reported in Table A8, the effect of the Austrian partition drops slightly, but remains significant and economically important. Depending on polynomial and bandwidth, the estimated coefficients range from 0.36σ to 0.45σ in the case of the 6th grade exam and from 0.2σ to 0.32σ in the case of the 9th grade. On the other hand, similarly as previously, the estimates of the effect of the Prussian partition are either not significant or have an opposite sign.

The current border between the vovoidships (NUTS2 administration level) overlaps almost completely with the former Russian-Austrian border. If vovoidships influence the quality of education, their effect could be mistakenly confounded with the effect of the Austrian or Russian partition. There are two arguments against this possibility. Firstly, the Polish education system is considered very decentralized (Herbst, Herczynski and Levitas, 2009). A local municipality's government manages the school network of almost all public elementary and lower secondary schools, at the same time the role of the central government is limited to financing education and enacting general resolutions. The vovoidship administration is thus practically irrelevant for the educational governance. Consistently, Table 6 shows that there is no significant effect of the Austrian Empire in municipality's educational spending per capita. Secondly, since the former Russian-Prussian border does not overlap with the administrative borders, I can include the vovoidship dummies in Equation (1). It turns out that none of these dummies are significant, which suggests that the vovoidship administration is not relevant for the performance of students. This observation is also consistent with Herbst (2004).

Finally, I run a set of placebo experiments in which step-by-step I artificially move the Austrian-Russian border by 5km to the North or to the South (at most around 100 km). I run the baseline specification, with the artificial borders, and define the "Russian" dummy as an area North from the artificial borders. Figure 9 Panel B presents Z-tests[32] of the placebo Partitions effects, for each artificial border. Notably, only the actual border (at point 0) is an outlier. Analogously, I move step-by-step the Prussian-Russian border by 5km to the West or to the East (at most around 100km) and define the "Russian" dummy as an area East from the artificial borders. Figure 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 Drwęca, 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

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

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

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 choose two artificial river borders with East-West orientation similar to the Austrian-Russian border. The first one goes along the Bug, the Vistula and the Bzura rivers; the second along the Wieprz, the Vistula and the Pilica rivers. Similarly, I select two artificial border with North-South orientation similar to the Prussian-Russian border. The first is located on the Narew, the Bug, the Vistula and the Pilica rivers. The second along the Odra river. In all cases, the effect of the placebo "Russian" dummy is 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, low and changes sign across specifications.

4 Channels of Persistence

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

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

4.1 Social Norms toward Local Education System

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

4.1.1 The Low Stake vs. The High Stake Exams

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

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

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

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

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

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

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

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

The optimal level of schooling effort is given by:

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

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

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

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

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

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

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

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

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

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

Which is consistent with the empirical results presented in Section 3, namely that the effect of the Austrian Empire is larger in the case of the Low Stake exam, than in the case of the High Stake one (Tables 5 - A8 and Figure 8). 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.

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

4.1.2 Proxies for Social Norms

I use survey data on attitudes toward education to see whether current social norms systematically differ across the historical borders of the Partitions of Poland. The primary source of the data is the two waves (2011 and 2013) of the Social Diagnosis survey, which include over 45,000 individuals from almost all counties in Poland. The data does not permit to directly measure social norms with respect to local schools. Nevertheless, the questionnaire asks, whether education is important for a good life, whether a respondents are satisfied with received education and their desired level of education for their children. The variables are described in Table A3. A second source, with smaller sample size, is the Life in Transition Survey (LiTS, organized by the European Bank of Reconstruction and Development), which includes around 7000 individuals from 350 primary sampling units (PSU) from Poland and asks questions about first or second priority of education in governmental spending, private expenditures on education and opinion about the role of intelligence and skills in life success. A third source is the Educational Value Added survey (EVA, conducted by the Educational Research Institute in Warsaw), which includes approximately 10,000 parents and asks about the role of family tradition in school selection. For the detailed descriptions of variables from LiTS and EVA see Table A2.

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

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

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

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

Table 7 (8) presents estimates of the effect of the Partitions of Poland using the sample from Social Diagnosis (LiTS and EVA) and the reported numbers are the marginal effects at the border (the average marginal effects). The results show that people living in the former Austrian partition are around 5% more likely to say that education is important for a good life (Table 7, columns (1) and (2)) and around 19% more likely to say that public education should be given first or second priority in governmental spending (Table 8, column (2), but (1) is insignificant). They are also 6% more likely to claim that family tradition is important in their local school choice ((Table 8, column (7), but (8) is insignificant). However, at the same time, they are around 20% *less* likely to desire higher education for their children (Table 7, columns (5) and (6)) and 20% *less* likely to agree that intelligence and skills are important in life success (Table 8, columns (3) and (4)). Finally, the sign of insignificant estimates suggests that people from the Austrian partition are more satisfied from received education ((Table 7, columns (3) and (4)), but spend less on education of their children ((Table 8, columns (5) and (6)).

These results can be interpreted as an evidence for a positive social norm toward the *local* educational institutions. People from the former Austrian Empire perceive education as important in their life, want more funds directed to public schools[36] and underscore the long run relationships of their families with local schools.[37] On the other hand, their attachment to the local schools might lead to a lesser trust toward the non-local ones. This would explain why people are less desirable to send their kids to an institution of higher education, which is usually outside the local environment. Nevertheless, they also perceive intelligence and skills as less important and, even though

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

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

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

this question is outside the educational context, this is a clearly puzzling result. Overall, in the light of the model outlined in the previous section, one could argue that people living in the former Austrian empire have higher social norms toward schooling effort in local school, that is $E(Austria) > E(Russia)$, but lower toward the expected future career and earnings $A(Austria) < A(Russia)$.

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

4.1.3 Kindergarten Attendance

The third piece of evidence shows that the Austrian partition has a positive and significant effect on kindergarten attendance. Table 9 presents the estimates of the border discontinuities in kindergarten attendance, defined as a share of children aged 3-5 who attend an institution of pre-education. I use the regressions specified as in Equation (1).[38] The results show that the kindergarten attendance ratio is higher on the Austrian side by 3-7 percentage points compare to Russia. On the other hand, there is no difference on the Prussian-Russian border.

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

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

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

4.2 Migration

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

Nevertheless, the effect might be affected by historical migration, if the selection of migrants is not orthogonal to the current performance of students. My main results could be, for instance, explained by migration of high skilled people from the Russian to the Austrian part (or low skilled vice versa) and/or migration of low skilled people from the Austrian part to third countries. Existing qualitative evidence suggests that the first possibility is unlikely. Labuda (1971) argues that majority of migrants between the partitions were seasonal workers and they did not settle permanently.[42] Moreover, the economic situation in Galicia was the hardest and the level of industrialization the lowest, so there were no strong incentives for skilled workers to migrate there. On the other hand, migration to third-countries is harder to assess, as there are no clear insights in the literature about the 19th century emigrants' skills from the partitioned Poland. Zubrzycki (1953) hypothesizes that migrants from the Grand Duchy of Poznań

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

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

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

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

were more likely to be better educated than migrants from Galicia and the Congress Kingdom.[43] However, as reported in Abramitzky, Boustan and Eriksson (2012), the late 19th century U.S. immigrants from Austria and Prussia had wages similar to the US population, even though there was a substantial variation in wages across immigrants coming from other European countries.

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

The post-World War 2 forced displacement of the Poles could possibly erase the effect of the Prussian Empire. The migration was from the eastern parts of the interwar Poland to the newly joined post-German areas on the western and northern of modern Poland. At the same time, the 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 a significant German minority before 1939, then the incoming Polish population could be 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 population and focus only on the areas with Polish majority. 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).[44] The upper part of Figure 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 the Germans were concentrated in the former Prussian partition, there is no visible discontinuity at the border with the Russian partition. The average share is

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

[44]Unfortunately, to my best knowledge, there is no census-level data available on the share of displaced people.The census of 1950 provides vovidship-level data. However, the vovidships on the Prussian side of the Prussian-Russian border include the areas joined to Poland after WW2 and thus the shares are naturally high there.

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 high share of Jews (see the lower part of Figure 10 and Table A9).

4.3 Other Channels

Alternative channels of persistence of the Partitions of Poland could be based on other differences between the Empires, such as urbanization and economic policies or the expansion of universities. In order to shed light on these, I estimate the border discontinuities of various socio-economic characteristics described in Table A1. I use the regressions specified as in Equation (1),[45] but with the dependent variables transformed to the natural logarithms. Note that, similarly as the exam scores, these variables are endogenous, and might reflect the effect of the historical education systems or some other channels.

Table 6 column (1) reports estimates of the effect (semi-elasticity) of the Russian Empire on the Prussian-Russian border, while column (2) on the Austrian-Russian border - therefore changing the sign yields either the effect of the Prussian or Austrian Empire. The effect of the Partitions on the Prussian-Russian border is insignificant, except for the share of people in agriculture and class size.

The picture looks different in the case of the Austrian-Russian border. Firstly, the Austrian Empire positively affects the demographic characteristics, namely the level of population, density, and migration balance. This could be an alternative channel if different urbanization patterns emerged during the Partitions period and urbanization influenced the quality of education. For instance, population density could affect class size, which in turn influences student performance (Angrist and Lavy, 1999).[46] However, there is no difference in class size between the former Austrian and Russian lands. Moreover, the reported estimations in Section 3.5 include the demographic characteristics as control variables. If urbanization patterns are the main channel of persistence, I should find an insignificant and small effect of the Austrian partition. Contrary to this, as reported in Table A8, the positive and strong effect of the Austrian partition is

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

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

insensitive to the inclusion of the demographic characteristics.

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

5 Identity as a Determinant of Persistence

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

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

5.1 Conceptualization

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

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

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

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

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

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

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

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

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

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

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

5.2 Evidence

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

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

I first document a simple correlation of the standardized 19th century educational attainment and the average of the modern exam scores (2005-2011). Figure 11 presents correlation for the 6th grade low stake exam score and Figure 12 for the the 9th grade high stake exam, in breakdown by the rural and total samples. Consistently with the hypothesis, we can observe positive correlations in the case of the Austria partition (which is significant for the 9th grade exam), but negative in the case of the Prussian partition (significant for the 6th grade exam and the rural sample).

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

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

where y_{mcwt} is the outcome variable for municipality m from county (deanery) c and partition p at time t (available for 2005-2011). A_{cp} is the educational attainment from the 1860s, which is available at the county (deanery) level, Aus_p is a dummy for the former Austrian partition, G_{mcp} is a set of exogenous geographic controls. In addition to this, in some specifications I include a set of time variant municipality socio-economic characteristics X_{mcpt}, defined in Table A1. I pool all the years and use the Random Effect estimator. The standard errors are clustered at the county (deanery) level.

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

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

Nevertheless, distinctive characteristics of the Austrian education system, other than the positive interaction with identity, might be another source of persistence. Especially important seem to be inclusiveness of the Austrian *secondary* education and the existence of two universities and one technical college in Galicia.[47]

6 Conclusions

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

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

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

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

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

The interaction between institutions and ethnic identity can be perceived as an ingredient of the inter-ethnic inequalities and modes of cooperation. As suggested by Jha (2013), a limited ethnic assimilation might lead to "the presence of nonreplicable and nonexpropriable source of 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 higher specialization in an economy, which, in turn, increases productivity.[48] In the context of Native American tribes, Dippel (2014) presents evidence that forced integration of linguistically homogeneous sub-tribal groups has a negative effect on the long run economic development through the quality of local governance. I add another channel to

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

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.[49]

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

GUS, "Drugi Powszechny Spis Ludności z dnia 9 grudnia 1931 roku. Mieszkania i Gospodarstwa Domowe. Ludnosć. Stosunki Zawodowe.," 1938.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

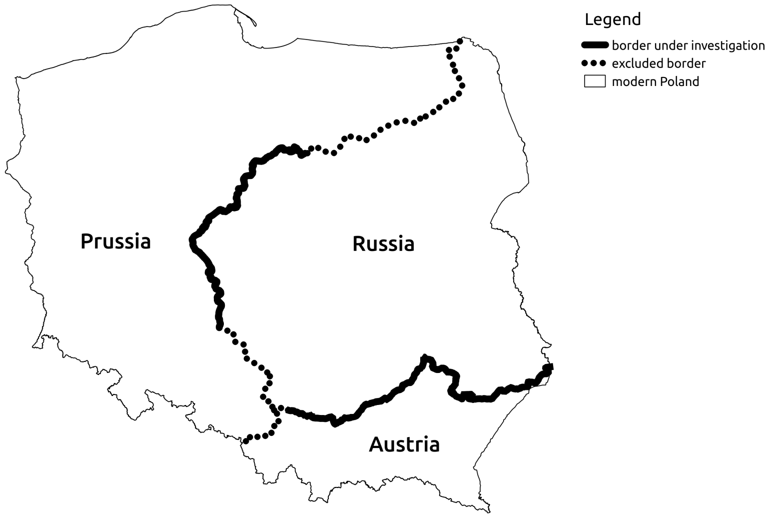
Sulimierski, Filip, Bronisław Chlebowski, and Władysław Walewski, *Słownik geograficzny Królestwa Polskiego i innych krajów słowiańskich*, Vol. 14, Filipa Sulimierskiego i Władysława Walewskiego, 1895.

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

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

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

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



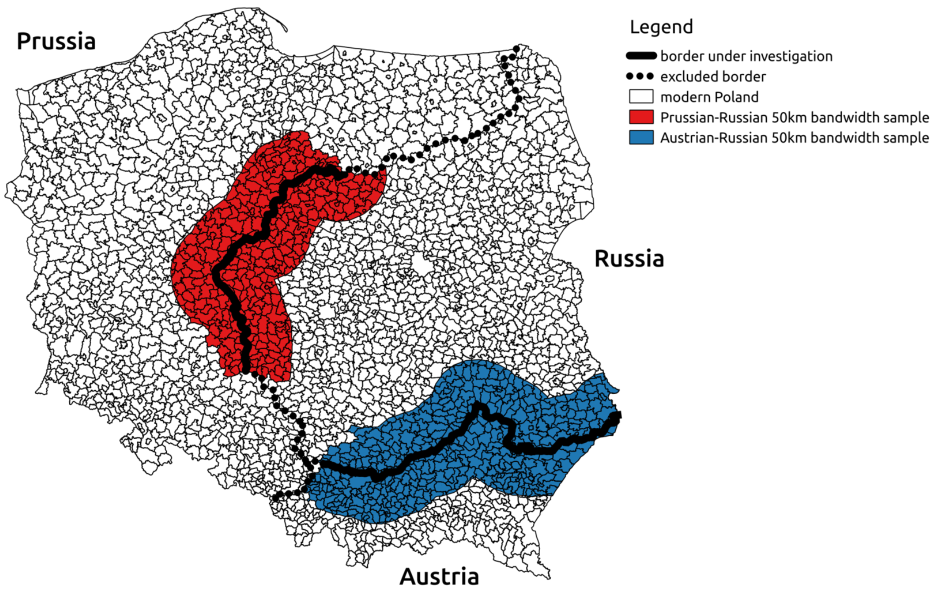


Figure 1: The Partitions of Poland 1815-1918

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

10.14754/CEU.2016.03

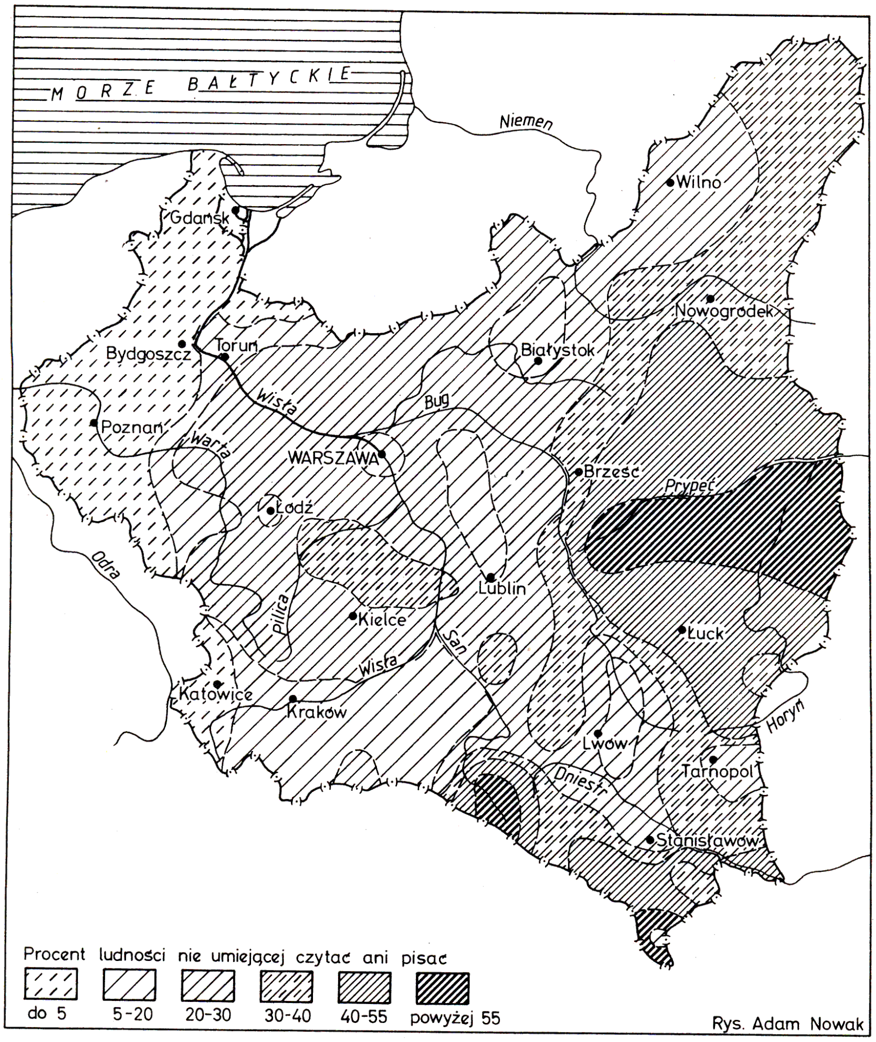


Figure 2: Illiteracy levels in the interwar Poland (1931)

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

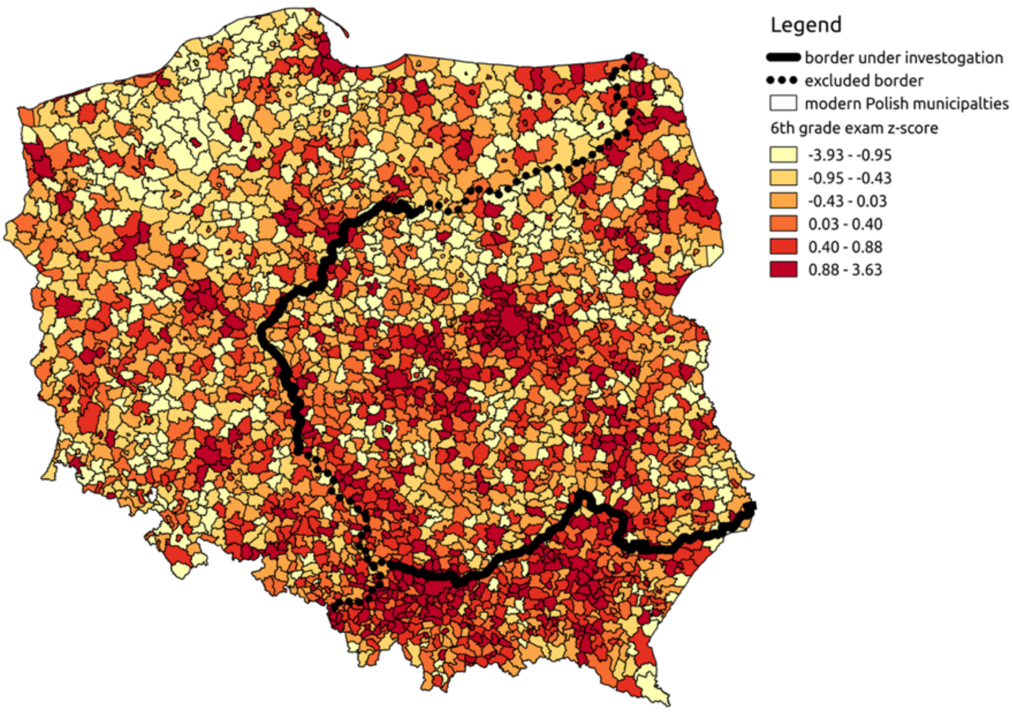


Figure 3: The distribution of the 6-th grade exam scores in 2011

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

10.14754/CEU.2016.03

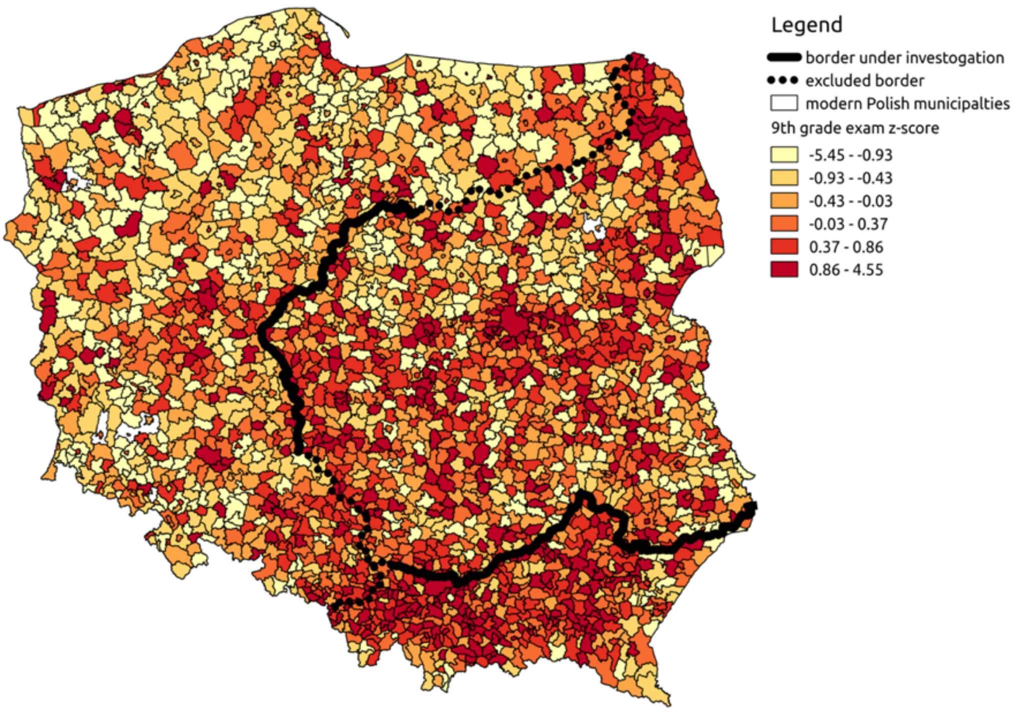


Figure 4: The distribution of the 9-th grade exam scores (math and science) in 2011

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

10.14754/CEU.2016.03

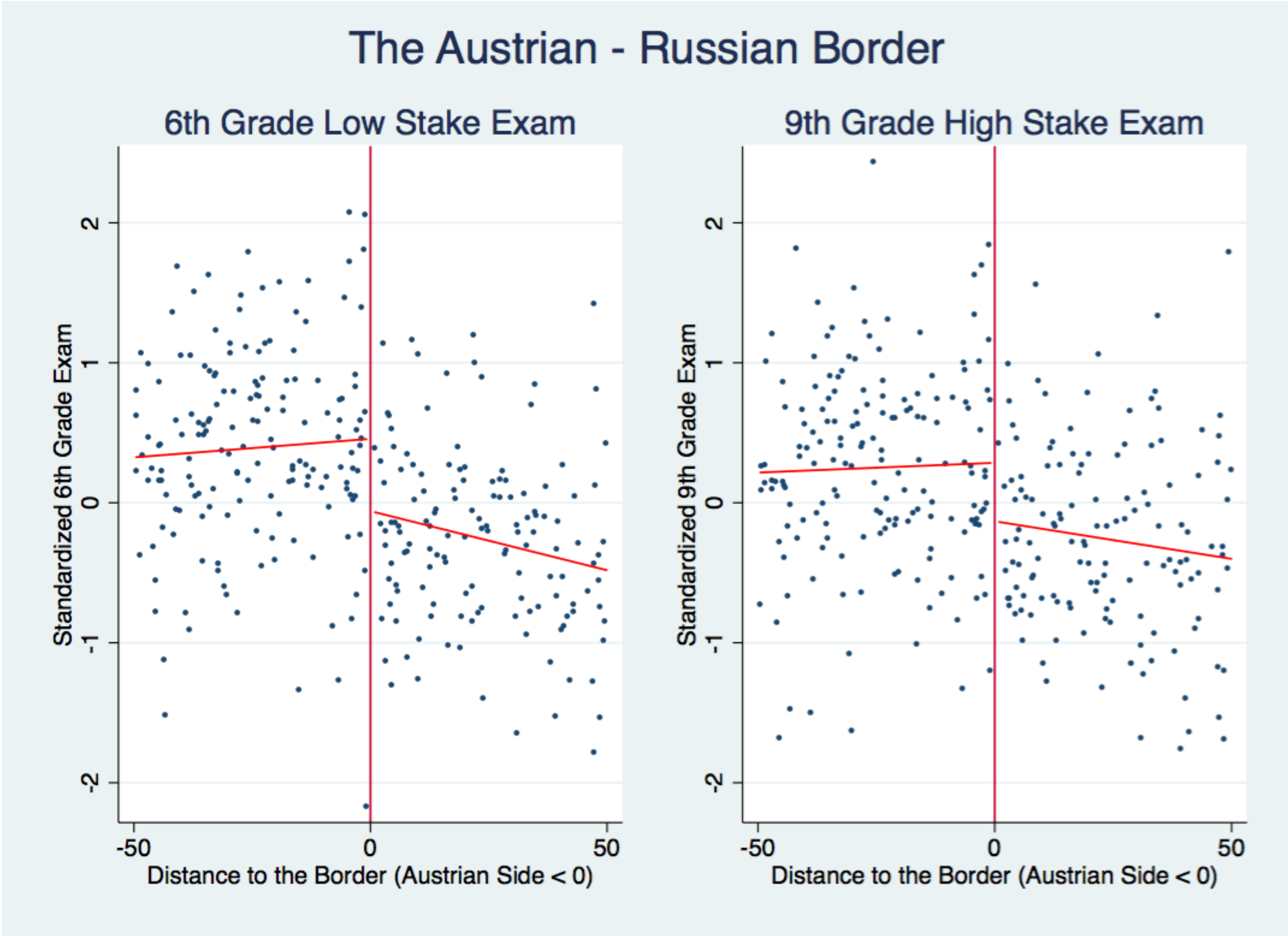


Figure 5: The one-dimensional regression discontinuity: Austrian-Russian border.

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

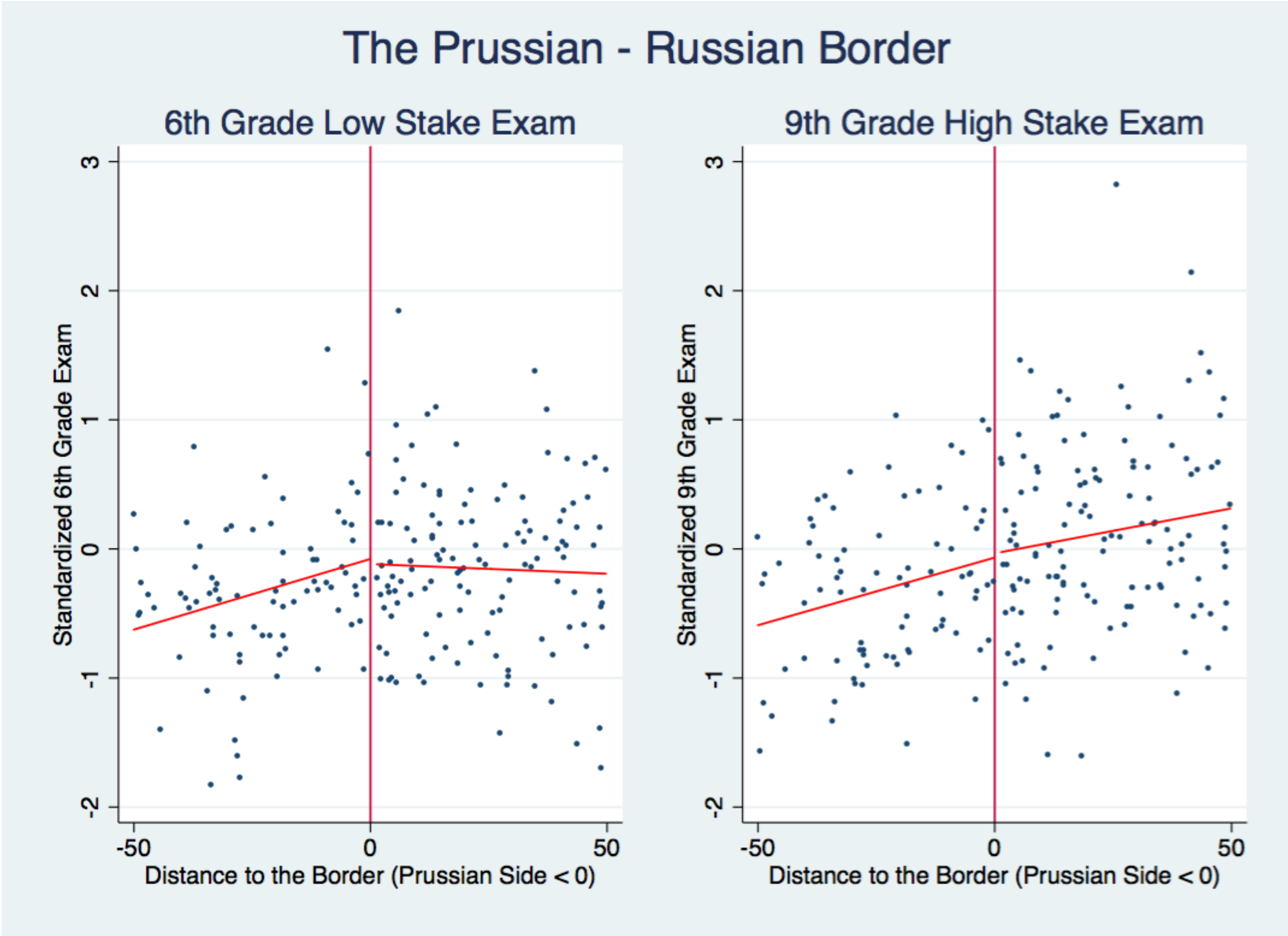


Figure 6: The one-dimensional regression discontinuity: Prussian-Russian border

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

10.14754/CEU.2016.03

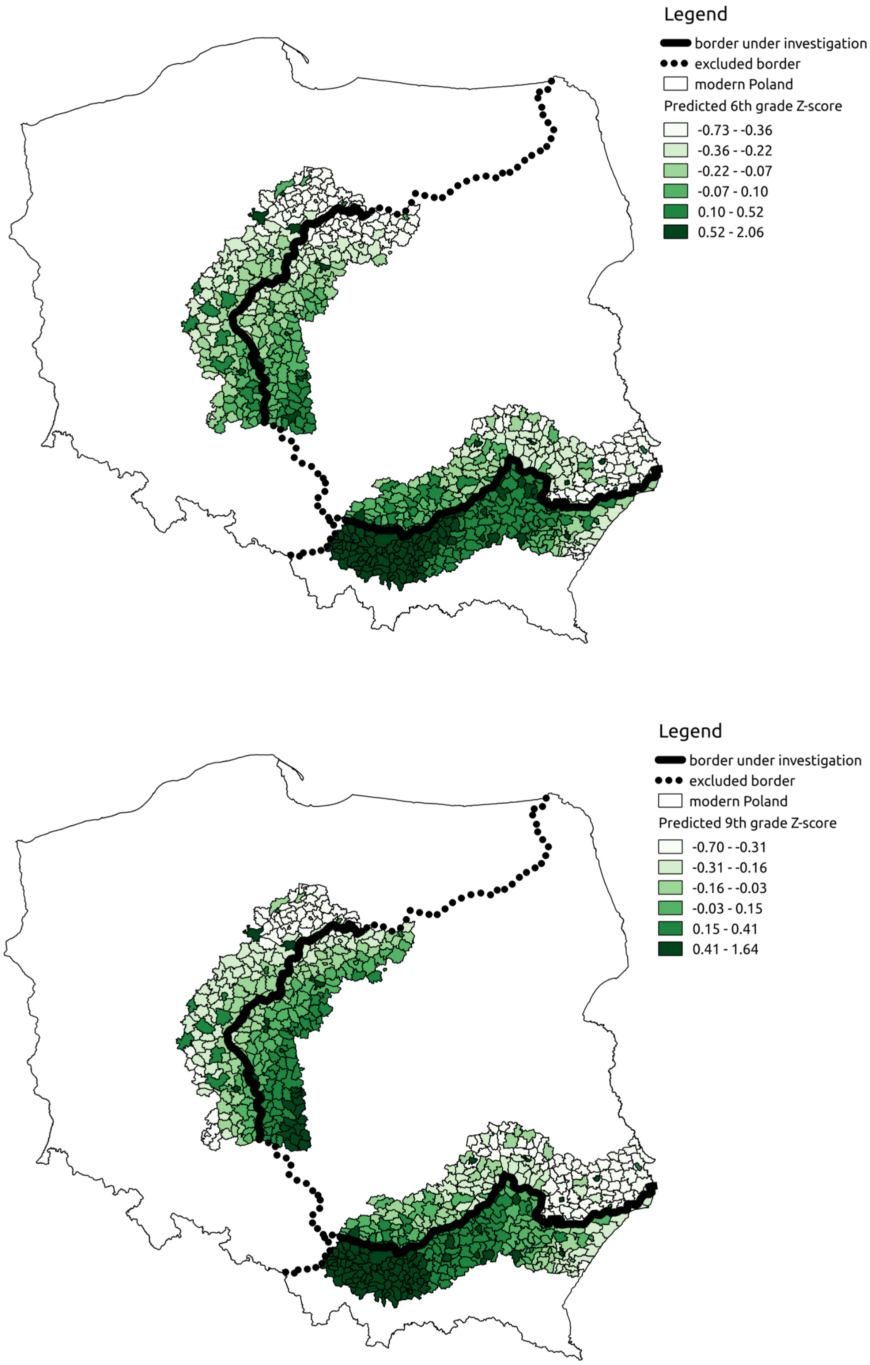


Figure 7: Predicted levels of the 6th and 9th grade exam Z-score for 2011

Note: the map shows the predicted values of the 6th and 9th grade exam Z-score for 2011, based on the regressions specified by Equation (1) for the whole sample, with the quadratic polynomial of longitude and latitude, the partition dummy, the geographic covariates and the set of population size dummies. The borders of the 19th century Partitions of Poland layered on the current map of Poland. The border under investigation are marked by the solid line, the excluded area by the dashed line. Silesia and Eastern Prussia were excluded because they belonged to Germany in the inter war period and were destination points for the massive post-WWII migration. Source: own visualization based on GIS maps from Kashin and Ziblatt (2012) and MPIDR and CGG (2012).

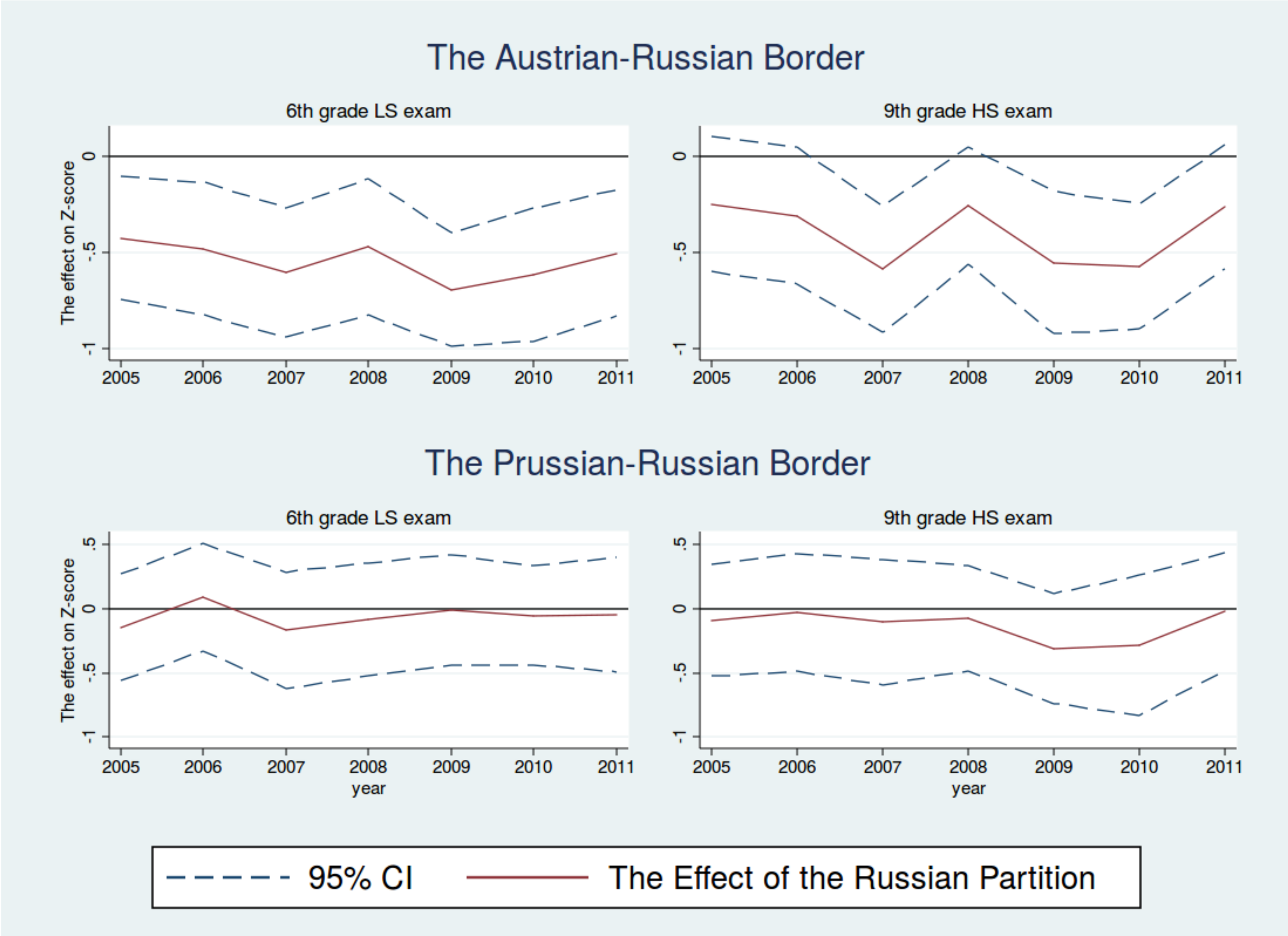


Figure 8: Estimations year-by-Y=year

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

10.14754/CEU.2016.03

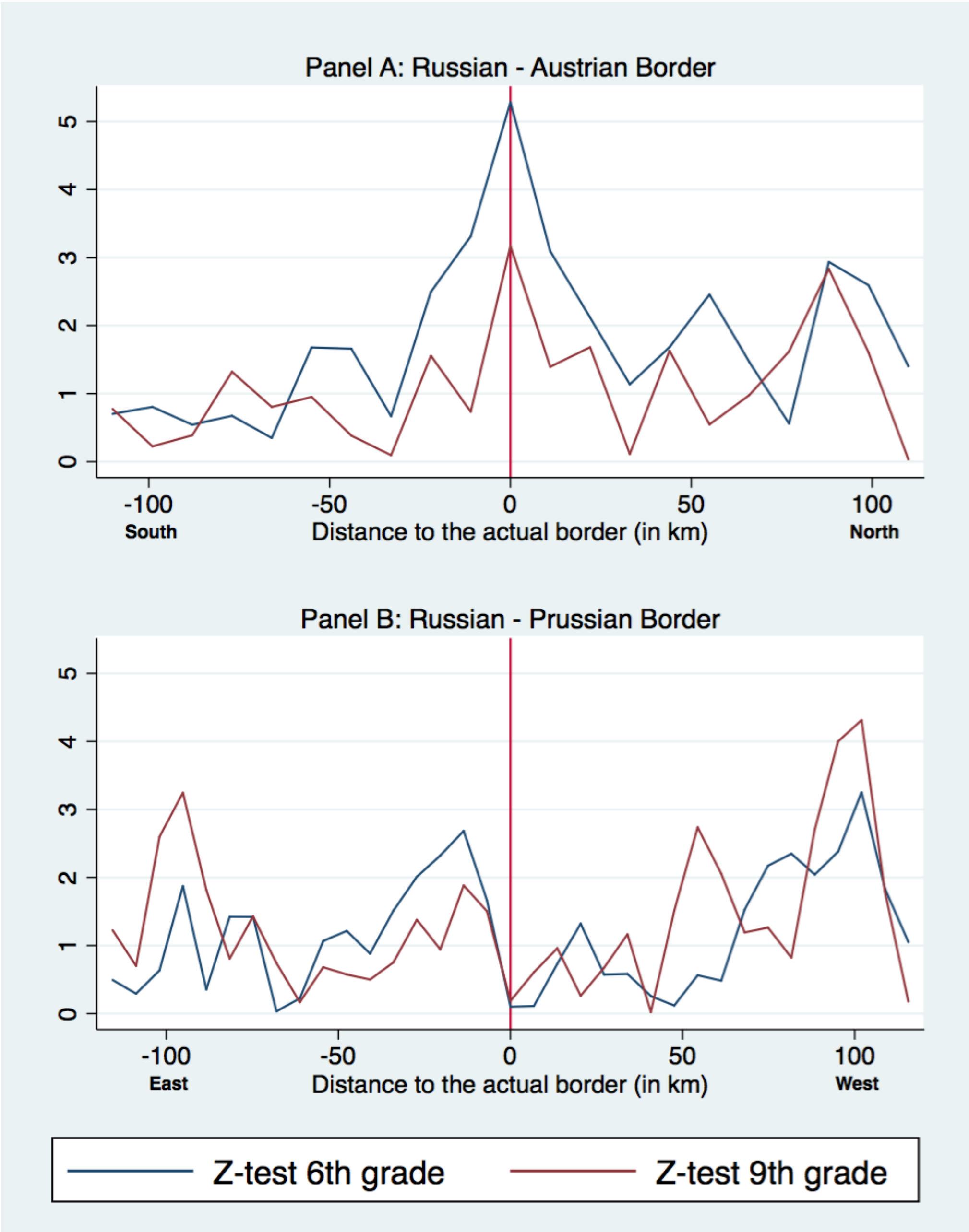


Figure 9: The placebo experiments

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

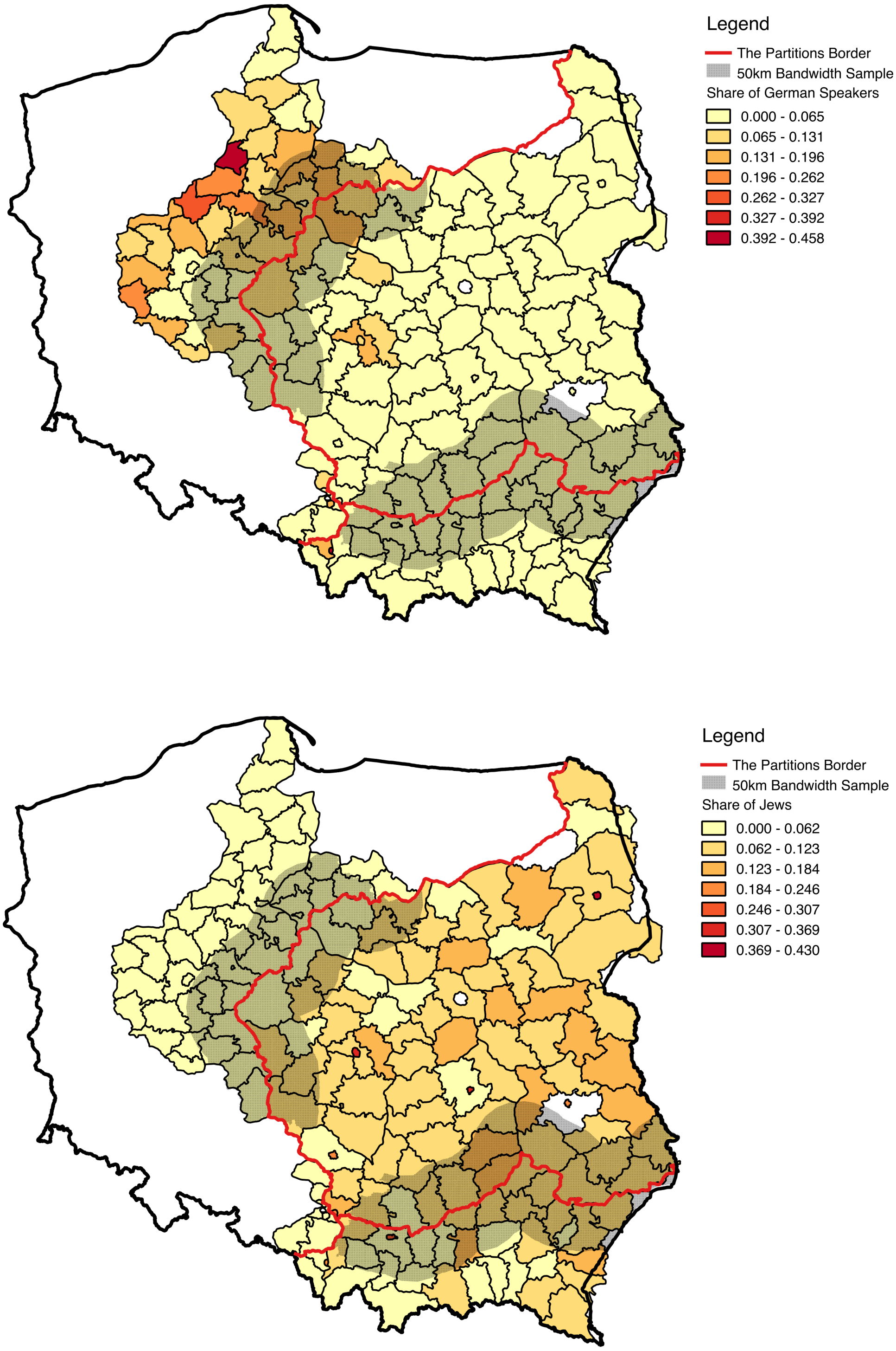


Figure 10: The Share of German Speakers and Jews in 1931

The map shows the 1931 counties, which were located on the territory of modern Poland. The upper map presents the share of German speaking population in the rural and urban areas. The lower map shows the share of Jews in the rural and urban areas.

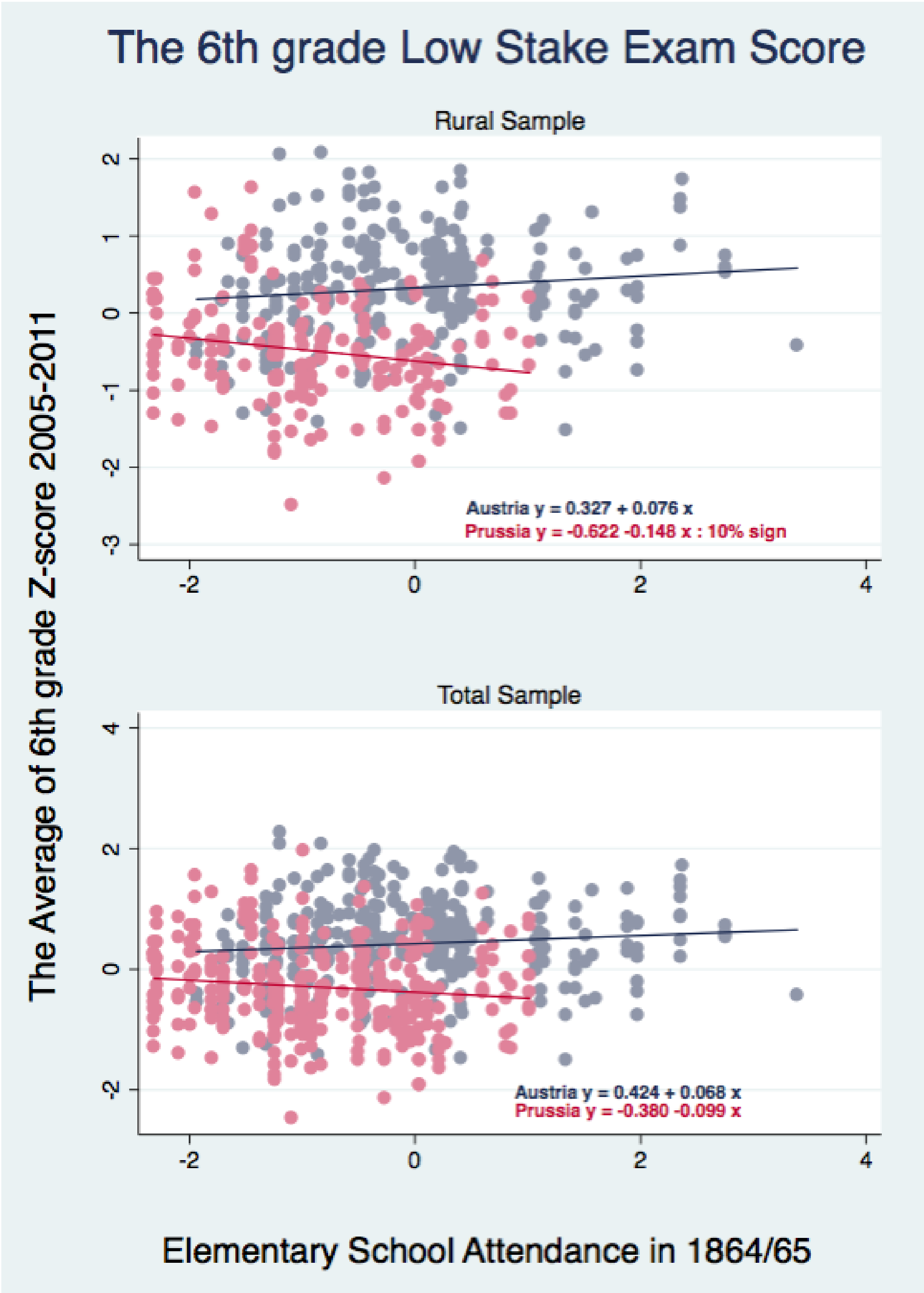


Figure 11: Historical educational attainment and modern performance of students

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

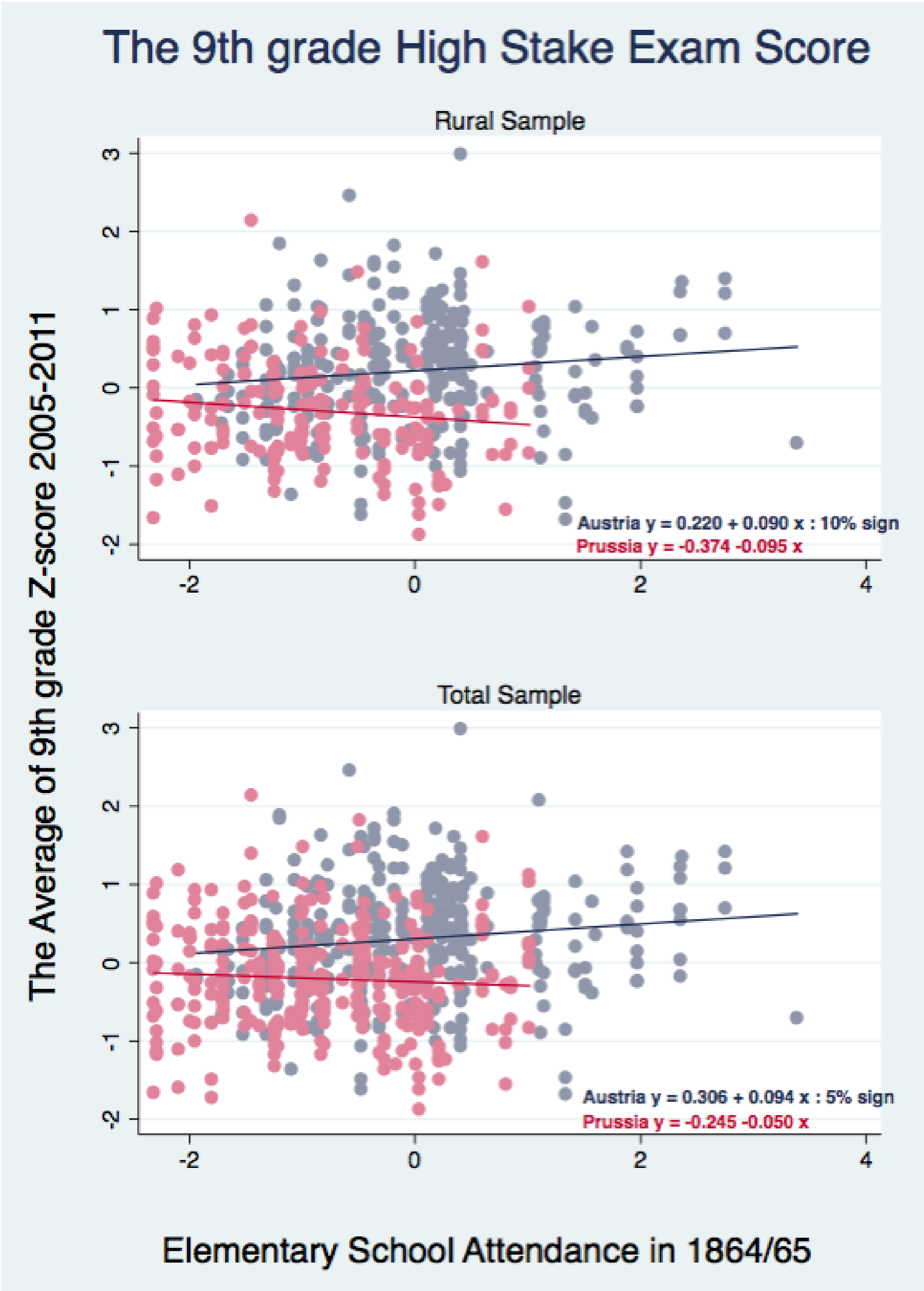


Figure 12: Historical educational attainment and modern performance of students

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

Table 1: Historical Characteristics of the Partitions

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

10.14754/CEU.2016.03

Table 2: Comparison of The 19th Century Educational Outcomes

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

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

Table 3: Descriptive Statistics

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

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

10.14754/CEU.2016.03

Table 4: Geographic Differences

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

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

Table 5: Baseline Regressions

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

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

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

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

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

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

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

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

10.14754/CEU.2016.03

10.14754/CEU.2016.03

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

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

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

Table 9: Kindergarten Attendance

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

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

10.14754/CEU.2016.03

Table 10: The Historical Data on Migration and Demographic Characteristics

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

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

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

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

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

10.14754/CEU.2016.03

Appendix

Table A1: Variables Description: The Regression Discontinuity Design

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

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

10.14754/CEU.2016.03

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.

10.14754/CEU.2016.03

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

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

Table A4: The Akaike Information Criteria

The Border:	Russian-Prussian		Russian-Austrian	
Outcome Exam:	6th grade	9th grade	6th grade	9th grade
	(1)	(2)	(3)	(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 a regression 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: Results: Polynomials in Latitude and Longitude

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

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

Table A6: Results: Polynomials in Latitude and Longitude, the Total Sample.

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

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

Table A7: Results: Polynomials in Distance

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

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

Table A8: Results: Polynomials in Latitude and Longitude, including Socio-Economic Covariates.

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

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

10.14754/CEU.2016.03

Table A9: Ethnic and Religious Composition in 1931

	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- all numbers are in percentage points										
Germans	7.172	4.656	6.28	5.153	.892	.255	.835	.089	.109	.166
Poles	91.031	4.484	86.833	3.72	4.198*	89.806	10.796	89.494	3.988	.311
Jews	1.804	3.219	6.63	2.771	-4.825**	8.274	4.862	10.736	4.177	-2.462
Greek-Catholics	.061	.039	.038	.008	.023	3.958	11.859	.041	.019	3.917
Orthodox	.112	.147	.099	.057	.013	3.417	9.405	.097	.126	3.32
Number of counties	29		6			26		10		
Panel B : Rural Areas <50km from the borders - all numbers are in percentage points										
Germans	8.631	5.467	7.543	5.481	1.089	.157	.406	.082	.128	.075
Poles	90.873	5.249	91.051	5.327	-.178	93.472	11.243	95.499	1.783	-2.027
Jews	.375	.869	1.065	.73	-.69+	3.099	2.244	4.584	1.786	-1.484+
Greek-Catholics	.05	.038	.035	.014	.015	4.412	12.999	.032	.011	4.38
Orthodox	.061	.096	.057	.028	.004	3.89	10.589	.042	.041	3.848
Number of counties	24		6			25		9		

Notes: Means and standard deviations for 1931's counties located at most 50km either from the former Russian-Prussian or Russian-Austrian border. Counties located in Śląskie vovoidships are excluded. Germans (Poles) is a share of Geman (Polish) speaking people in population (total or rural). Jews, Greek-Catholics and Orthodox are analogous but consider religion affiliation, not language. ** denotes significance at the 1% level, * at the 5% level and + at the 10% level.

10.14754/CEU.2016.03

Chapter 2
The threat of competition and public school performance: evidence from Poland *

joint with Martyna Kobus

Theoretical literature on whether competition from private schools raises public school productivity is ambiguous (e.g. MacLeod and Urquiola, 2015) and empirical literature is scarce (e.g. Hsieh and Urquiola, 2006). We provide evidence for the negative effect of the *threat* of competition on students' test scores in public schools in Poland, which has a decentralised educational system and experienced large improvements in international student exams (PISA). The identification strategy uses the introduction of the amendment facilitating the creation of autonomous schools in Poland in 2009 as an external shock to the threat of competition. We focus on the short run in which there is only a limited set of actions available to schools' principals. For the total sample we find no effect, however, for more competitive urban educational markets, we report a drop in test scores in public schools following increased threat of competition. This negative effect is robust to the existence of autonomous schools prior to the amendment and to the size of public schools. It does not result from a pre-existing or concurrent trend either. We exclude student sorting and school's expenditures adjustments as potential channels.

*We thank Sascha O. Becker, Roman Dolata, Torberg Falch, David Figlio, Jan Herczyński, Gábor Kézdi, Małgorzata Kłobuszewska, Sergey Lychagin, Aneta Sobotka, Mateusz Żółtak ,the participants of seminars at Central European University and The Educational Research Institute in Warsaw and The Workshop on Educational Governance in Trondheim, the 6th RGS Doctoral Conference in Economics in Bohum and EALE Conference in Warsaw for their comments and suggestions. Martyna Kobus' participation in the project was supported by the Polish Ministry of Science and Higher Education under a "Mobility Research Award" (contract No. Nr 893/MOB/2012/0). All opinions expressed are those of the author and have not been endorsed by the Ministry. All errors are ours.

10.14754/CEU.2016.03

1 Introduction

Disconnection between educational expenditures and student achievement (Hanushek, Mayer and Peterson, 1999) as well as between standard measures of school quality and student achievement (Hanushek, 2003) have turned economists' attention to the incentive structure of public schools (Betts, 1995; Hoxby, 2003). There is a substantial disagreement in the literature on how market-like incentives should impact on public school performance. For those who argue that following increased competition public schools should improve their quality (Hoxby, 2003), the basic argument is that more productive schools would drive away students from their current school. This process would continue until higher quality schools dominate the whole educational market or public schools respond to competitive pressure. This would be true if (i) private schools were indeed more productive (which is a separate empirical question), and if (ii) public schools reacted to competition by increasing productivity. Some authors McMillan (2005), MacLeod and Urquiola (2013) argue that the latter does not have to be the case. Schools may find it optimal to exert lower effort if the losses from the smaller market share are more than offset by the saving in effort cost. McMillan (2005) argues that rent-seeking public schools may find it optimal to reduce productivity following increased competition. In communities in which, they decide to exert lower effort and let the high-income type students leave. MacLeod and Urquiola (2013) analyse the literature on the the economics of industrial organisation, contract theory and asymmetric information and conclude that this literature suggests that there is no a priori reason to believe that greater school choice results in higher school productivity. In other words, the theoretical literature does not provide a clear sign for the productivity changes in public schools due to increased competition. Empirical evidence is also very scarce and consists of 3-4 papers. It suffers from identification problems i.e. ideally an exogenous variation in the size of competition is needed. Furthermore, it is difficult to isolate productivity changes from other processes such as changes in the student composition and in peer effects. The available evidence concentrates mostly on large scale voucher reforms in Chile and Sweden and exploits changes in private enrolment. This controls for factors that are constant over time i.e. parental background.

In this paper we extend empirical literature by providing evidence on the negative effects of competition. The evidence suggests that theoretical model of school competition may be incomplete in the short run. In particular, we analyse impact of the threat of competition from community schools on public school performance in Poland.[1] Community schools are more autonomous than public schools with respect to teachers' hiring and salaries and collecting external funds, however, they have to follow a nationwide curriculum. As an identification strategy we use the amendment to the education act introduced in March 2009 which facilitated the creation of community schools but only for schools that have 71 and less students. Therefore, public schools located in areas where there is a higher percentage of students who attend schools with less than 71 students are more exposed to competition. We show that this is indeed strongly related to actual creations of community schools. Using year 2009 as a breakthrough date in the Difference-in-Difference estimations we find that the higher threat of competition caused by the mentioned reform has significant negative impact on the performance of urban public schools. Urban areas are more competitive educational markets than rural areas, with dense school network and better parental background. The effect is similar for public schools that are larger (more than 300 students) and becomes stronger for urban schools that already have a community school in their neighbourhood and may thus be more aware of the consequences of the reform. The generalized model and placebo experiments show it does not result from a pre-existing or concurrent trend either. We argue that in our case changes in students composition between schools are unlikely, but also similarly to Hsieh and Urquiola (2006) we analyse changes in test scores at the municipality level, which cancels out the direct effects of sorting between schools. In cities (above 20000 inhabitants) test scores drop at the aggregate community level.

Our paper is most similar to Hoxby (2003, p.32) in that there is an exogenous variation in the *threat* of competition, which we exploit. As states Hoxby (2003, p.32) "choice options do not arise randomly, but are frequently a response to school conduct". Consequently, the *actual* change in competition might be endogenous with respect to the characteristics of educational markets. Therefore, it is the threat of competition that matters. Comparing to Hoxby (2003, p.32) we do not discreticize our treatment variable. We focus on the short run consequences of the reform. In the short run, decisions regarding the school might be different than predicted by reputational model of (MacLeod and Urquiola, 2009). If school's reputation depends on both productivity and on the ability of its students, private schools have the incentive to boost their reputation by cream skimming the best students rather than by raising productivity. Then, increasing choice may result in lower average school productivity. This is consistent with what happened

[1]Poland has experienced substantial gains in PISA scores, rising from 470 points in 2000 to 518 in 2012, placing Poland fifth in Europe and eleventh in the world.

in Chile. In 1981 Chile introduced nationwide school choice by providing vouchers to students, resulting in 20 percentage points increase in private enrollment rate. Hsieh and Urquiola (2003) find that the main effect of this expansion is school stratification. Comparing the change in educational outcomes in urban and wealthier communities to that in communities where private schooling increased by less the authors find no evidence that more competition improved test scores, but that repetition and grades actually worsened. On the contrary, if schools cannot select on ability (e.g. they must select students via lotteries), then their model implies that school choice will unambiguously raise school performance and student outcomes. It seems that this is what happened in Sweden. In 1991 Sweden introduced a reform which led to the creation of independent municipality fund schools and to greater school choice as students were allowed to attend any school in their jurisdiction. Admissions into elementary level were based on proximity to the school, so cream-skimming was not allowed. Bohlmark and Lindhal (2012) analyse all ninth-grade cohorts from 1988 to 2009 and report an increase in a broad set of outcomes following the increase in competition e.g. test scores in language and math, grades in language and math, university attendance and average years of schooling at age 24. Sandstrom and Bergstrom (2005) analyse how higher share of independent schools impacts on the students results in the neighbouring public schools. To account for the endogeneity in the share of independent schools, they use proxies for the general attitude that a given municipality has towards delegating its duties to the private sector. Municipalities more open to the private sector are less likely to block independent schools entry into the market. They find that greater school competition leads to better performance of public school students on standardised exams.

In the short run, school adjustment to increased competition may include activities that are attractive and visible to parents (e.g. school trips), but are not necessarily related to productivity. As Hoxby (2003) states, in the short run principals "(...) can induce their staff to work harder; they can get rid of unproductive staff and programs; they can allocate resources away from non-achievement-oriented activities (building self-esteem) and toward achievement-oriented ones (math, reading, and so on)". In Poland, teachers enjoy high level of employment security (ensured by the collective bargaining arrangement, so-called *Karta Nauczyciela*) and they cannot be laid off easily, in particular, not in the short run. Therefore, what remains available to school principals, is either efficiency changes i.e. incentivising teachers to work harder, or boosting their school's prestige. We find no evidence for increasing teachers' salaries or investment in infrastructure in the expenditure dataset. Anecdotal evidence suggest that principals may deter the entry of community schools by shifting their attention from tasks oriented at performance of students, to those which are visible to parents (e.g. school trips).

Polish community schools differ from such schools in other countries, but in general community schools have their idiosyncratic aspects (Heers, Van Klaveren, Groot and van den Brink, 2011). In Poland such schools are not targeted at low or high-income students, they operate like regular schools, however they are given a substantial autonomy when it comes to the management. As such it makes them also similar to charter schools in the USA or autonomy schools in the UK. The main reason for their creation is, however, financial. The move toward greater cost rationalisation of the Polish local governments and the actions taken by the central Polish government to reduce fiscal debt force many *gminas* (i.e. municipality - NUTS 5 administration unit) to close small elementary schools and move students into bigger establishments. Parents object to the closure of a school because of the increased distance to a new school and the role the school plays in cultural and social life of the local community. Thus closures lead to significant tensions between local governments and citizens. One way to release these tensions is to allow non-public associations (i.e. typically parents' associations) to take control over a school that is planned to be closed, and create non-public community school instead. Local government provides the community school with a subsidy financed by the central government (i.e. tied to the number of pupils), but it does not need to cover the remaining costs which can be as high as 50 percent. Community schools are free of charge and are more autonomous than public schools. In particular, they do not need to abide by collective bargaining agreements with teachers' unions that determine teachers' salaries. They are also more flexible in acquiring external funds, but they have to follow a standard, nation-wide curriculum.

The threat of handover of a small school to a non-public association creates competitive pressure for other local public schools for two reasons. Firstly, since central government subsidy is tied to the student, a public school loses money when a student goes to a community school. For the same reason, liquidation of a small school is beneficial for another public school, because students from a liquidated school attend the public school. Secondly, because of the cost-effectiveness of community schools, local politicians may consider the transformation of all public schools into community schools. Principal-agent problems make it unlikely that principals form a coalition with parents and transform their school into a community school. Principals are more free in their decisions when they respond to a local politician than directly to parents whose kids attend the school. The principals are motivated to influence local politicians to close endangered schools and thus block the entry of new schools. Therefore, the ensuing number of liquidations and community schools is endogenous and subject to the degree of competition.

This paper contributes to the literature on more autonomous schools too, although we do not study such schools themselves, but rather their spillover effects. Clark (2009) analyses British reform that allowed public high schools gain more autonomy. He finds no evidence of spillover effects of such schools on the schools in their neighborhood. Machin and Vernoit (2010) study the introduction of academy schools in UK that allowed schools to change their structure and become more autonomous. They report positive spillovers to neighboring schools. There is also related literature on the effects of decentralization in the US (Hoxby, 2000; Rothstein, 2007). This, however, provides choice between public school districts rather than between private and public schools, so this literature answers a different question and so far has produced mixed results. Similarly for the effects of private voucher-induced competition on public school performance in the USA (Hoxby, 2002; Figlio and Hart, 2014; Abdulkadiroglu, Angrist, Hull and Pathak, 2014).

The paper is organised as follows. In Section 2 we describe the educational system in Poland with particular emphasis on community schools. In Section 3 we present empirical strategy and data. Section 4 contains the results and a series of robustness checks. In Section 6 we discuss the results and conclude.

2 Community schools in Poland

2.1 Polish educational system and community schools: brief overview

The Polish comprehensive education is compulsory and consists of six years of elementary school (ISCED 1), which is followed by three years of lower secondary school *gimnazjum* (ISCED 2). Elementary school and *gimnazjum* usually serve the same community of students, but they are separated entities, with different managerial and teaching bodies. After finishing the comprehensive part, student may finish their education or continue in academic, mixed or vocational higher secondary school (ICED 3).

The admission to elementary school and *gimnazjum* is identical. It is based on catchment areas, which means that every student has a right to attend an assigned local public school and this school has to accept her. As an alternative to the local school, parents may request a place in an under-subscribed non-local school, but the admission is not granted. The are no universal recruitment rules for non-local students. Each school's policy is determined by school principal and a recruitment committee, which usually consists of selected teachers and school psychologist.

During the comprehensive education, students are examined by two standardized, externally graded and obligatory examinations: a low stake exam after elementary school (6th grade) and a high stake exam after *gimnazjum* (9th grade).[2] The school averages from these exams are published in various unofficial school rankings. The only official measure of school quality is the school-level educational value added, but it is only available for *gimnazja*.[3]

Polish education system (and Polish administrative system in general) is considered to be very decentralized (Herbst, Herczyński and Levitas, 2009). This can be clearly seen in the distribution of power among different levels of government. All public elementary and secondary schools (*gimnazja*) are governed by local governments, and the role of central government is limited to education financing and imposing a core curriculum. Local governments are free to open and close new schools, hire teachers, principals and redistribute money among schools.

There are clear economies of scale for school principals. All Polish public schools and some non-public schools are financed by the central government through subsidies. In theory their amount should be sufficient to cover all expenditures on education (excluding investments and pre-school education). In practice, however, it covers only around 50-70% of the costs (Instytut Badań Edukacyjnych, 2011; Herbst et al., 2009) and the rest is covered by local governments. Since the governmental subsidy is tied with the pupil (the money goes with her), the amount of money which is given to a specific school depends on its enrollment. Furthermore, principals might gain more local political power from larger schools, which might be crucial in securing additional funds from the municipality government.

The teacher wages and general employment conditions are mostly determined by municipality governments (not by school principals) in compliance with the universal collective bargaining agreements (*Karta Nauczyciela*). Salary has to be at least as large as a minimum wage determined for each teacher's rank in *Karta Nauczyciela*.[4] In addition,

[2]The 9th grade exam serves as a basis for the admission into the higher secondary education (ISCED 3).

[3]It is widely available only since 2009. The Ministry of Education publishes also annual level-based rankings of the higher secondary school.

[4]In 2015 the minimum monthly gross wages ranged from 1513 PLN (340 EUR) to 3109 (700 EUR). Additionally, local municipalities have to make sure that the average *total* gross salary for each tacher's rank within municipality is at least as large as specified in *Karta Nauczciela*. In 2015 these

teachers may receive extra salary for working over-time, monetary awards and other non-monetary benefits e.g. accommodation on school's premises.

Because of fixed costs, a smaller school yields higher cost per student. Thus, greater cost-rationalisation creates incentive for local governments to merge schools and close them down. However, such decision is politically dangerous for a local government, therefore community schools appear as an attractive option. Such schools are led by non-public associations and are also financed from the public money, but the local government does not have to finance expenditures that exceed the amount of the subsidy. Community schools are more autonomous than public schools, but they have to follow a nation-wide curriculum. In particular, teachers, who are employed in regular public schools, sign the contract with their corresponding local government but the lower bound of their wages is set by the central authority (in agreement with the teachers' unions). Moreover, they have several benefits e.g. the right to a year-long sick leave, which is exceptional among occupational groups financed from the central budget. Conversely, the specifics of the contract with the community school are negotiated at the school level and the wage is regulated by the free market. Since the wages are almost 80% of the educational spendings, community schools can manage the expenditures and the composition of a teaching body much more flexibly. It is also important that community schools have much better opportunities than public schools in terms of financing of new investments and operational costs from external funds, such as European Union funds. Finally, parental involvement in community schools is much higher than in public schools since community schools are often led by parents' associations, so the principal - agent problem is less severe.

2.2 The amendment

Before 2009 a school could be taken over by the association only after it had been liquidated first i.e. a new school had to be set up. Liquidation was a complicated process and many activists and politicians saw the need for an improvement of this procedure. The possible solution was to allow a takeover without putting the school into liquidation. The first sign of possible change appeared in 2007 when PO (Civic Platform), a major political party supporting the idea of a community school, won the parliamentary election. Half a year later, around June 2008 the first official project of the amendment was created, and it immediately ignited public debates. Within the next

averages ranged from 2717 PLN (612 EUR) to 5000 PLN (1126 EUR).

half a year the opposition was trying to block the reform by saying that it woulds lead to the privatisation of public education. The teacher unions in December 2008 organised the nation-wide campaign against schools' handover "Do not let our school get ruined". After almost a year of ongoing debate, the amendment to the education act 1991 was finally introduced in March 2009. It allows the takeover of schools without putting them into liquidation[5] by a natural or legal person other than a public authority, when the school's enrolment is smaller than 70 pupils. This number has no specific meaning. Figure 1 shows the distribution of school size in Poland; it is "smooth" around the value 70. The mean size is 192 and the median size is 125. The distribution is relatively dispersed, standard deviation is 172.

The main remaining problem with schools' handover is that in order to obtain the governmental subsidy for the first four months of operation, a non-public association has to inform the authority about the planned take-over around 11 months before the first day of operation. If an association does not manage to do this, it is up to a local government to act in their stead. Although the amendment did not completely resolve the legal problems faced by associations and the new amendments are being discussed, 2009 amendment is considered a very important step toward the full introduction of community schools into the Polish education system.

There are currently 11398 public elementary schools and 949 schools run by associations. Not all such schools were a result of hand-overs, but this is now the main channel through which schools run by associations are established. In years 2008-2012 446 schools were handed over to non-public associations. Indeed, the amendment seems to have facilitated handovers, as they accelerated significantly after its introduction. While in 2009 there were 30 handovers, in 2011 the number was 89 and in 2012 - 244. In year 2010 only 19 handovers took place, but this is due to elections. Historical data (since 2002) confirm the hypothesis that election years are characterised by very small changes in the school network (Herczynski and Sobotka, 2013). Whereas before 2009 local governments relied mostly on liquidation, in year 2012 handovers became as common as liquidations. Usually elementary schools are transformed into community schools (84,9% of all handovers). Most handovers happen in eastern Poland which is less densely populated so that there is higher percentage of schools with enrolment less than 70 students.

[5]The old way of a hand-over through liquidation was left as an option, though.

10.14754/CEU.2016.03

3 Measuring the competitive effect of the threat of handover

3.1 Empirical strategy

As we mentioned, we focus on how the threat of competition affects the performance of public schools i.e. student achievement in external exams. The threat of competition is measured as the fraction of students in a given municipality who attend a school with 70 or less pupils. The amendment facilitated the handing over of such schools only. The existence of small schools is likely to be correlated with unobservable characteristics of *gminas* and therefore simple regression of outcome on the proportion of small schools might be biased. Yet, as long as these characteristics do not change over time, one can control for them by focusing on changes. We exploit the introduction of the amendment to the Education Act 1991 from March 2009 as a breakthrough point, and Difference-in-Difference methodology to overcome identification difficulties. In other words, we claim that schools that have low exposure to treatment i.e. almost unchanged threat of competition following the reform, act as a control group for schools that face a bigger threat. If schools with low exposure faced in fact bigger threats, then our results can be seen as a lower bound (i.e. in absolute terms). We use a set of controls, in particular we control for the population size of the municipality which determines the structure of school network to some extent. Our outcome variable is the results of standardised and obligatory nation-wide external exams taken by students at the end of elementary school. We look only at the performance of schools larger than 70 students. Our goal is to capture the effect of threat of competition, not the threat of being transformed into a community school.

We follow Card (1992) and use the panel fixed effect estimator to estimate the following model

$$S_{git} = \beta_(T_g \times After) + \delta X_{gt} + \mu_g + \mu_t + \mu_{gi} + \epsilon_{git} \tag{1}$$

where S_{git} is average test score in public elementary school i located in municipality g at time t, T_g is fraction of students in municipality g who attend to elementary public school with less than 70 pupils in year 2008. Higher value of this variable means bigger threat of competition after the introduction of 2009 amendment. The time dummy $After$ switches on for observations following the introduction of the amendment i.e. years 2009-2012. “Before” means period 2005-2008. μ_t are year-specific effects, X_{gt} are time-variant characteristics of *gminas*, μ_g and μ_{gi} are unobserved time-invariant municipality's and school's characteristics. Finally, ϵ_{git} is the error term.

The parameter of interest is β. Main hypothesis is that it is positive i.e. we expect that higher threat of competition induces public schools to improve their outcomes or that it causes an inflow of good students from schools under a threat of handover. The latter effect is, however, unlikely in our case (see Section 4.2). Moreover, one should expect that the effect is heterogeneous in different sub-samples. In particular, there are substantial differences in Poland between rural and urban educational institutions (Jakubowski and Kozińska-Bałdyga, 2005). We run separate analyses for both urban and rural areas.

3.2 Data

Our sample consists of 9846 publicly funded elementary schools with enrolment above 70 students (robustness checks restrict the sample to bigger entities).[6] The data on the results of the exam at the end of an elementary school (ISCED 1) from years 2005-2011 come from Centralna Komisja Egzaminacyjna (the Central Examination Commission). This is a database which contains results of all exams in all schools in Poland. The exam score is generally irrelevant for further education[7], however, it is obligatory and is considered by the local authority in the school evaluation process. Students are examined in reading, writing skills, logical reasoning and the usage of knowledge in solving more practical problems. The maximum score is 40 points. The exam is not standardised across years, although we standardise it by subtracting the mean and dividing by standard deviation each year.

The treatment variable requires information on school enrolment in 2008. This comes from System Informacji Oświatowej (the System of Educational Information), the registry database about all schools in Poland. It also contains rich set of school characteristics including a panel of school spending (on renovations and teacher salaries), school infrastructure (green and sport areas) and changes in school network (school handovers and closures). These data are available for 9192 publicly funded elementary schools (larger than 70 students) for years 2007-2011. Data on the characteristics of a gmina for years 2005-2011 comes from Central Statistical Office of Poland.

[6]Typical elementary school in Poland has six grades and children start education at the age of 7.

[7]The secondary school (*gimnazjum*) is obligatory and there is a legal obligation to accept a pupil to a school which corresponds to where he or she lives. However, some non-public secondary schools might consider the exam results in the recruiting process.

We could obtain information on school size for only 70% of observations. For the rest we approximate the number of students by multiplying the number of students who took the test by the number of grades (i.e., six).

4 Results

4.1 Descriptive statistics

Table 2 presents the municipality level descriptive statistics for 2008. In general rural areas in Poland are significantly poorer and less developed than urban areas. There are also important differences in the structure of educational market, especially in the share of students studying in schools larger than 70 - the treatment variable. On average 9.2% (s.d. 12.3pp) of students attend small schools, but this number varies across the rural and urban areas. In the former it is slightly higher than 10% (s.d. 12.6pp) , whereas in the latter it is only 1% (s.d. 1.7pp). The maximum share for an urban municipality is 11.1 %, for a rural over 78%. Interestingly, 4.7% of schools in the urban areas are led by communities, while in the total and rural samples it is 3%.

Table 3 presents cross correlations between the treatment variable and municipality characteristics in 2008. It shows that in the urban areas, the higher the exposure to the reform, the lower the secondary school gross enrolment ratio and the higher the expenditures on education per capita. However, the magnitudes are not economically significant. In the total and rural samples, the treatment variable is positively correlated with unemployment rate and educational expenditures, and negatively with total expenditures, population level and density, kindergarten and secondary school enrolment and with number of students. Overall, the higher the fraction of smaller schools in the area, the worse the municipality's characteristics. This is partially explained by the fact that there are more small schools in lesser populated ares, which are also generally poorer.

4.2 Main results

First, we motivate our claim that the reform increased the threat of competition. Table 4 Columns (1) and (2) show the effect of the higher exposure to the reform on the probability of school handover. The model is similar to (1), except that the dependent variable is at municipality level and takes value one if there was an episode of school handover in municipality-year and zero otherwise. The data is available only for 2007-2011. The results show that 10pp increase in the treatment variable leads to around

1pp increase in the probability of school handover. However, the effect is heterogeneous, in the urban sample it is ten times larger, i.e., 10pp increase in the treatment variable leads to 10pp increase in the probability of handover. These results show that the 2009 reform, especially in the urban areas, had a significant effect on the local educational market as it increased a probability of entering a new type of school and thus decreased probability of small school liquidation.

The first column of Table 5 Panel A shows main results from the panel fixed effect estimations of the baseline model, controlling for the year-specific effects but without the additional covariates. Because the estimator is exploiting the variation from within an observation, the unobservable and observable time - invariant characteristics of schools and *gminas* cancel out. The impact of the variable of interest is negative but insignificant. It becomes larger in absolute terms when one adds educational covariates such as municipality's gross enrolment in pre-school education, secondary education ratio and expenditures on education per capita. Ten percentage point increase in the fraction of small schools in the area causes a drop in the exam score on average by 0.009σ of exam score. The effect is similar once we add covariates describing the general economic condition of a municipality: population size, unemployment rate and total expenditures per capita. We do not have complete data for all *gminas*, therefore 8 schools are dropped from the regression with the educational covariates. Neither do we have a balanced panel for some of these covariates.

As shown in Table 4, the effect of the reform was more likely to be visible in the urban areas. To check for the heterogeneity in outcomes we run regression 1 for rural and urban samples separately. There are 7144 rural schools and 2702 urban schools in our dataset. The first two columns of Table 5 Panel B report estimations for the rural sub-sample, with and without covariates. Results show that the parameter of interest is much smaller and statistically not different from zero, however, for the urban sub-sample (Panel C) the effect is, in absolute value, several times bigger than in the baseline model and significant at 1% level. A standard deviation increase in the treatment variable causes a drop in the exam score on average by 0.026σ of exam score. The change in the value of treatment from minimum to maximum causes a decrease in exam score by 0.26σ. Adding covariates does not change the magnitude of the results (Table 6).

These results suggest that the introduction of the amendment causes changes in school performance only among urban schools and the decrease in test scores is relatively large. This can be explained by the characteristics of the educational market which is larger in urban areas, more competitive (i.e. dense school network makes schools more accessible) and characterised by higher demand due to e.g. better educated and thus more motivated parents. To be more precise, population density in the urban areas equals 1656 persons per square kilometre versus 166 in the rural area; the ratio of elementary schools per square kilometres equals 0.21 in the urban area versus 0.028 in the rural area; for public transport network the numbers are 1.656 vs. 0.071. Finally, there is only 4% of tertiary education degree holders in rural areas comparing to 11% in urban areas.

4.3 Heterogeneity and robustness

We run four additional analyses. We take into account that the results might be different for schools that already have a community school in their neighbourhood and for larger schools (more than 150 and 300 students). We also study the results of the secondary school exam, which acts as a placebo test. Secondary school handovers are extremely rare, thus the amendment was unlikely to affect operation of these schools. Finally, we also test for the common trend assumption by running a generalised version of (1), with interactions between the treatment variable and each year dummy.

The key assumption in the identification strategy is that *gminas* with different intensities of treatment have the same pre-treatment trends in outcomes. A possible source of heterogeneity in the time trend may come from the pre-treatment existence of community schools among schools which were exposed to the reform. Since there are potential spillovers of community schools on the performance of other schools, it might be the case that if the performance had been falling prior to the reform, we would overstate the negative impact of the reform. Also public schools which have a community school in their vicinity are more aware of the impact of such schools on their situation, and of the potential consequences of the reform. Therefore, we limit our sample only to public elementary schools located in *gminas* where there was at least one community school in year 2008. There are now 2731 schools in our dataset. Table 7 shows results for the restricted sample and rural and urban sub-samples.Comparing to the baseline scenario, the parameter of interest is larger and becomes significant in a specification with the full set of covariates (column (4)). The existence of a community school potentially makes public schools more aware of the consequences of the reform and induces a stronger effect. For urban schools, the magnitude of the results is similar to previous results. A standard deviation increase in the fraction of small schools in 2008 causes a drop in the exam score on average by 0.031σ of exam score.

Our main sample consists of public elementary schools with enrolment higher than 70 students in 2008. Given the heated debate about the amendment before its official introduction and the fact that the threshold was chosen rather ad hoc, there is a possibility that schools with enrolment just above 70 students were facing the probability of handover. This might cause specific reaction of these schools to the introduction of the amendment. Since there is a negative correlation in our sample between a school size and the treatment intensity, our results can be driven by this effect. We limit the sample only to public elementary schools that have more than 150 students which is close the median size. Furthermore, Jakubowski (2006) argues that marginal cost of additional students becomes flat for schools bigger than 150 students. This might suggest that closures of such schools may be less profitable for the local governments. Table 8 reports the results for sample of 2227 (1570) urban and 2725 (854) rural schools larger than 150 (300) students. In case of schools that have over 150 students the effect is insignificant in the whole sample and negative and significant in urban areas. The magnitude is slightly smaller than in the unrestricted urban sample. For schools over 300 students in the total sample, the effect is marginally insignificant and twice larger than in unrestricted set of schools, but in the case of the urban areas the effect is weaker.

We check the impact of the reform on the exam taken at the end of secondary school. This is a sort of placebo test, because the amendment should not cause any activity on the part of secondary schools. We find that in the urban sample, the impact of the treatment on secondary school exam results is insignificant, which suggests that the negative impact that we find for elementary schools is not associated with any concurrent trends that are underway in the municipality. Recall that for the whole sample of elementary schools, we find no effect. For the whole sample of secondary schools we find significant and positive impact. Ten percentage point increase in the fraction of students learning in schools up to 70 students causes an increase in test scores by 0.019σ (Table 9).

Finally, in order to test the existence of pre-existing trends we run a generalised version of (1), with interactions between the treatment variable and each year dummy:

$$S_{git} = \sum_{t=2006}^{2011} \beta_t (T_g \times \mu_t) + \delta X_{gt} + \mu_g + \mu_t + \mu_{gi} + \epsilon_{git} \qquad (2)$$

The notation is similar as previously (see Section 3.1.) Note that year 2005 is a reference point, thus for instance β_{2006} is the effect of the exposure to the treatment in 2006 relative to 2005.

The reform was introduced in 2009, therefore if the common trend assumption is true only interactions after 2009 should be significantly different from zero. Figure 2 plots the β_t coefficients against time (x-axis) along with confidence intervals, for the urban sub-sample only. It shows that indeed only years after 2009 are significantly different from zero. In addition, the F-test for the joint significance of the interaction terms prior to 2009 does not allow to reject the null hypothesis that the effects are zero.

5 Channels

5.1 Sorting of Students

One explanation of the negative result are the flows/self-selection of students between schools i.e. sorting (Hsieh and Urquiola, 2006). This has a direct effect of composition of students on the results and an indirect effect through peer groups. Yet in our case selection of students is unlikely. For the negative result to hold, either good students would need to go from public school to a community school or bad students from a community school would need to move to a larger public school. These are, however, at best medium - run consequences of the reform whereas we observe negative effects in the short run.

Another possibility is that the reform increased probability of small school closure. This is against the reform, which was aimed at saving small school, but could be a result of strategic behaviour of large school principals (discussed in the next subsection). If students from small schools are on average under-performers, then the small school closure and relocation of students to larger schools, might explain the reported negative effect. However, Table 4 Columns (3) and (4) show that, while there is a significant effect of the reform in the total and rural samples, it is insignificant in the urban sample. This is inconsistent with the negative effect of the threat of competition among urban schools and strongly suggest that the changes around the extensive margin i.e. the school closure channel, seem to be unlikely.

Finally, we follow Hsieh and Urquiola (2006) and measure aggregate change in outcomes at the municipality level. This strategy at least controls for the direct effect of sorting i.e. student composition, but is unable to net out the peer effects. If the documented effect is due to the flow of students within municipality, there should be no effect on the aggregated measure. Table 10 presents the results. We find that the effect is insignificant in total population (Panel A) and it is negative and significant among urban schools (Panel B). Maximal change in the treatment variable causes a 0.2σ drop in the exam score.

5.2 Strategic Behaviour of School Principals

One explanation of the negative effect of the higher exposure to the amendment on the performance of public school is that principals react to bigger competition threat by allocating resources in a way that is disadvantageous for student achievement measured by test score. Principals want to "advertise" their schools to parents, local politicians and local public opinion. *System Informacji Oświatowej* provides school-level data on the share of spending devoted to renovations and teacher salaries, and on the size of green and sport areas per student. In order to see whether the higher threat of competition, induced by the reform, had any effect on these outcomes, we run (1) with the above-mentioned measures as dependent variables. Table 11 reports the results. We find no evidence that the higher threat of competition leads to changes in the school-level spending on renovations, teacher and on the size of green and sport areas. These suggest that either this channel is not important or principals' attempt to "advertise" their schools is not fully captured by these variables.

There is some anecdotal evidence supporting this mechanism. The report about social tensions emerging during a handover of a school, published by Ośrodek Rozwoju Edukacji (Centre for Educational Development) documents reaction of a large public elementary's principal from Kożuchowo. For example, in Kożuchowo (Kloc, 2012) a local small school was planned to be liquidated, however local parental organisations opposed it and proposed the handover of the school to a non-public association. The principal of the large public school (which would take over students in case of liquidation) was trying to persuade parents that his school is better than the newly established community school: "[...] for the parents of students [from the small school] he organised attractive and competitive curriculum. Additional sport and language classes, after-class activities, school trips, cafeteria, bus transportation of students, safety were the main points. This offer was passed over to all parents." (Kloc, 2012, p.19).

Principals can also devote time and resources to influencing politicians' preferences through lobbying. Local governments have several options to consider when the financial situation forces them to reorganise the school network. Generally the most effective solution from the financial perspective is liquidation. The municipality does not have to pay costs of buildings and moving students to a bigger entity decreases the average cost of student in this institution (because of decreasing marginal costs). Yet once one considers the political costs of liquidation, namely voters' protests, the balance of gains and losses can change. Such protests against schools' closures are visible in local and nation-wide media. Two other possible moves are to phase out a small school or to liquidate only part of grades e.g. usually grades from 4th to 6th (in case of an elementary school) are closed and students are moved to another school. Both scenarios are more costly than the liquidation and only a bit less controversial for the local community. Therefore the last solution - school handover - is considered as a realistic alternative to the liquidation. Yet the municipality still has to practically take care of the school building[8] and provide the governmental subsidy.

In general, the attitude of local governments toward community schools is ambiguous. On the one hand, some local governments realise that handing schools to associations, even when they are not under the threat of liquidation, helps to ease financial tensions. An anecdotal example of this attitude might be the voyt of Hanna in the eastern Poland who gave all public schools in his municipality to parental associations (Grabek, 2013). This is, however, a rather extreme case. On the other hand, local political leaders often have close relationship with public schools' principals. Also, teachers are usually a strong and well-organised interest group at the local level (Kloc, 2012). Both of these factors act against the establishment of community schools. A similar scenario took place in Bielany in the eastern Poland where the voyt and one of the principals were trying to block a small school handover. They did not succeed because the municipality council finally supported the parental association.

6 Conclusions

We provide evidence that competitive effects caused by the threat of establishing more autonomous schools in the area are negative, significant, and appear mainly in urban areas. Polish reform provides for a credible proxy for the threat of competition and we focus on it. This allows us to account for the fact that actual entries into educational markets may be endogenous to school conduct or reactions on the part of public schools' principals. In Polish case, principals have the incentive to block the entry of a community school. We focus on the short run. As implied by the reputation model of MacLeod and Urquiola (2009), in the absence of cream-skimming, more competition will unambiguously lead to better school performance and student outcomes. Cream-skimming is, however, unlikely in our case and we do observe a robust negative effect of higher competitive pressures. It might be the case that when the decisions are made in the short run, apart from established school's reputation based on its effectiveness and student

[8]The local governments are usually renting a building to the parental association for a symbolic price, like 1 PLN.

composition, simple marketing mechanisms aimed at attracting parents may be at play and resources may be used to finance them. Then measures such as promotion of the accountability system among parents and local governors might decrease the negative impact of the competition.

References

Abdulkadiroglu, Atila, Joshua D. Angrist, Peter D. Hull, and Parag A. Pathak, "Charters Without Lotteries: Testing Takeovers in New Orleans and Boston," Working Paper 20792, National Bureau of Economic Research December 2014.

Betts, Julian R, "Does school quality matter? Evidence from the National Longitudinal Survey of Youth," *The Review of Economics and Statistics*, 1995, pp. 231–250.

Bohlmark, A. and M. Lindhal, "Independent schools and long-run educational outcomes: Evidence from Sweden's large scale voucher reform," IZA Working Paper No. 6683 2012.

Card, David, "Using regional variation in wages to measure the effects of the federal minimum wage," Technical Report, National Bureau of Economic Research 1992.

Clark, Damon, "The performance and competitive effects of school autonomy," *Journal of Political Economy*, 2009, *117* (4), 745–783.

Figlio, David and Cassandra Hart, "Competitive effects of means-tested school vouchers," *American Economic Journal: Applied Economics*, 2014, *6* (1), 133–156.

Grabek, Artur, "Oswiata bez karty nauczyciela," 2013. Online; posted 20-December-2013, http://www.rp.pl/artykul/19,1074030-Oswiata-bez-Karty-nauczyciela.html?p=2.

Hanushek, Eric A, Susan E Mayer, and Paul Peterson, "The evidence on class size," *Earning and learning: How schools matter*, 1999, pp. 131–168.

Hanushek, Eric Alan, *The Economics of schooling and school quality*, Vol. 159, Edward Elgar Pub., 2003.

Heers, Marieke, Chris Van Klaveren, Wim Groot, and H Maassen van den Brink, "Community schools unfolded: A review of the literature," *TIER*, 2011.

Herbst, Mikołaj, Jan Herczyński, and Anthony Levitas, *Finansowanie oświaty w Polsce: diagnoza, dylematy, możliwości*, Wydawnictwo Naukowe Scholar, 2009.

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

Hoxby, Caroline M, "Does competition among public schools benefit students and taxpayers?," *American Economic Review*, 2000, pp. 1209–1238.

_ , *How school choice affects the achievement of public school students*, Vol. Choice with Equity, Hoover Press, Stanford, 2002.

Hoxby, Caroline Minter, "School choice and school productivity. Could school choice be a tide that lifts all boats?," in "The economics of school choice," University of Chicago Press, 2003, pp. 287–342.

Hsieh, Chang-Tai and Miguel Urquiola, "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of public Economics*, 2006, *90* (8), 1477–1503.

Hsieh, C.T and M. Urquiola, "When schools compete, how do they compete? An assessment of Chile's nationwide school voucher program.," National Bureau of Economic Research Working Paper No. 1008 2003.

Instytut Badań Edukacyjnych, "Raport o stanie edukacji 2010," 2011.

Jakubowski, M and A Kozińska-Bałdyga, "„Małe Szkoły–polska odpowiedź na wybrane problemy wiejskiej oświaty"," *Wieś i Rolnictwo*, 2005, *3*, 2005.

Jakubowski, Maciej, "Efektywność instytucji zarządzających lokalnymi dobrami publicznymi na przykładzie inicjatywy Mała Szkoła." PhD dissertation, WNE Uniwersytet Warszawski 2006.

Kloc, Kazimierz, "Konflikty w procesie racjonalizacji sieci szkół," Technical Report, Ośrodek Rozwoju Edukacji 2012.

Machin, Stephen and James Vernoit, "The Effects of Moving to More Autonomous School Structures: Academy Schools and Their Introduction to English Education," *Centre for Economic Performance, London School of Economics*, 2010.

MacLeod, W. B. and M. Urquiola, *Competition and educational productivity: incentives writ large*, Vol. Education Policy in Developing Countries, University of Chicago Press, 2013.

MacLeod, W.B. and M. Urquiola, "Anti-lemons: school reputation and educational quality," National Bureau of Economic Research Working Paper No. 15112 2009.

_ **and** _ , "Reputation and school competition," *American Economic Review*, 2015, *105* (11), 3471–3488.

McMillan, P., "Competition, incentives, and public school productivity," *Journal of Public Economics*, 2005, *89*, 1131–1154.

Rothstein, Jesse, "Does Competition Among Public Schools Benefit Students and Taxpayers? Comment," *American Economic Review*, 2007, *97* (5), 2026–2037.

Sandstrom, M. and F. Bergstrom, "School vouchers in practice: competition won't hurt you," *Journal of Public Economics*, 2005, *89*, 351–380.

Table 1: Descriptive Statistics (2008)

	N	Mean	SD	Min	Max
		Unit of Observation: Municipality			
Panel A : Total					
Standardized 6th grade exam	2477	0	1	-3.349	3.378
Unemployment rate	2477	.076	.038	.009	.249
Expenditures per capita	2477	2547.362	1066.835	1606.25	45560.129
Population(th.)	2477	15.395	50.653	1.346	1709.781
Share of population 0-18	2477	21.262	2.48	11.068	32.673
Population density (th. per km^2)	2477	.244	1.28	.005	59.528
Educational Expenditures per capita	2477	929.321	183.434	454.04	2846.18
Kindergarten Attendance	2477	.511	.184	0	1.203
Sec. School gross enrollment ratio	2475	.965	.171	0	2.058
Number of elementary school students	2477	964.692	2276.638	79	70248
Share of students in schools >71	2477	.094	.123	0	.788
Share of community schools	2477	.032	.091	0	.8
Panel B : Rural subsample					
Standardized 6th grade exam	2250	-.075	.995	-3.349	3.378
Unemployment rate	2250	.077	.038	.009	.249
Expenditures per capita	2250	2521.068	1078.252	1606.25	45560.129
Population(th.)	2250	8.558	5.503	1.346	50.161
Share of population 0-18	2250	21.554	2.348	11.068	32.673
Population density (th. per km^2)	2250	.107	.195	.005	2.883
Educational Expenditures per capita	2250	931.018	177.636	454.04	2846.18
Kindergarten Attendance	2250	.489	.175	0	1.203
Sec. School gross enrollment ratio	2248	.951	.167	0	2.058
Number of elementary school students	2250	617.665	397.291	79	3546
Share of students in schools >71	2250	.102	.126	0	.788
Share of community schools	2250	.03	.093	0	.8
Panel C : Urban subsample					
Standardized 6th grade exam	227	.745	.71	-1.361	2.729
Unemployment rate	227	.06	.026	.011	.136
Expenditures per capita	227	2807.99	907.933	1785.78	8662.73
Population(th.)	227	83.159	150.765	17.897	1709.781
Share of population 0-18	227	18.37	1.806	12.3	23.022
Population density (th. per km^2)	227	1.595	3.944	.076	59.528
Educational Expenditures per capita	227	912.503	233.064	558.97	1554.27
Kindergarten Attendance	227	.732	.098	.064	.989
Sec. School gross enrollment ratio	227	1.109	.14	.154	1.651
Number of elementary school students	227	4404.383	6490.844	803	70248
Share of students in schools >71	227	.009	.017	0	.111
Share of community schools	227	.047	.074	0	.357

Table 2: Descriptive Statistics (2008) - municipalities with at least one community school in 2008

	N	Mean	SD	Min	Max
		Unit of Observation: Municipality			
Panel A : Total					
Standardized 6th grade exam	356	.266	1.007	-2.994	3.378
Unemployment rate	356	.071	.036	.009	.249
Expenditures per capita	356	2613.257	638.132	1606.25	6160.66
Population(th.)	356	42.281	125.93	1.704	1709.781
Share of population 0-18	356	20.515	2.826	14.02	30.112
Population density (th. per km^2)	356	.571	3.214	.006	59.528
Educational Expenditures per capita	356	946.962	191.544	454.04	1682.08
Kindergarten Attendance	356	.56	.195	.108	1.126
Sec. School gross enrollment ratio	356	.986	.154	0	1.972
Number of elementary school students	356	2227.43	5512.801	93	70248
Share of students in schools >71	356	.092	.121	0	.567
Share of community schools	356	.221	.128	.022	.8
Panel B : Rural subsample					
Standardized 6th grade exam	272	.047	.985	-2.994	3.378
Unemployment rate	272	.076	.038	.009	.249
Expenditures per capita	272	2475.028	492.771	1606.25	6082.6
Population(th.)	272	9.943	6.186	1.704	42.976
Share of population 0-18	272	21.405	2.494	14.606	30.112
Population density (th. per km^2)	272	.09	.149	.006	1.522
Educational Expenditures per capita	272	933.185	169.649	454.04	1682.08
Kindergarten Attendance	272	.497	.175	.108	1.126
Sec. School gross enrollment ratio	272	.948	.148	0	1.972
Number of elementary school students	272	670.066	427.719	93	2952
Share of students in schools >71	272	.116	.129	0	.567
Share of community schools	272	.249	.128	.063	.8
Panel C : Urban subsample					
Standardized 6th grade exam	84	.977	.711	-.739	2.729
Unemployment rate	84	.054	.024	.014	.122
Expenditures per capita	84	3060.856	826.73	1929.88	6160.66
Population(th.)	84	146.996	230.606	22.665	1709.781
Share of population 0-18	84	17.631	1.678	14.02	21.437
Population density (th. per km^2)	84	2.126	6.397	.131	59.528
Educational Expenditures per capita	84	991.574	245.595	628.71	1523.35
Kindergarten Attendance	84	.762	.089	.535	.984
Sec. School gross enrollment ratio	84	1.108	.103	.933	1.651
Number of elementary school students	84	7270.321	9782.781	1239	70248
Share of students in schools >71	84	.012	.018	0	.111
Share of community schools	84	.128	.069	.022	.357

Table 3: Descriptive Statistics: Correlations with the Treatment Variable (2008)

	Dependent Variable: Share of students in schools >71		
Sample:	Total	Rural	Urban
	Unit of Observation: Municipality (2008)		
Standardized 6th grade exam	-.010 (.002)***	-.004 (.003)	-.0005 (.002)
Unemployment rate	.262 (.065)***	.181 (.069)***	-.063 (.044)
Expenditures per capita	-9.20e-06 (2.30e-06)***	-7.90e-06 (2.50e-06)***	1.10e-06 (1.20e-06)
Population(th.)	-.0003 (.00005)***	-.005 (.0005)***	6.80e-06 (7.40e-06)
Share of population 0-18	.004 (.001)***	.0001 (.001)	-.0006 (.0006)
Population density (th. per km^2)	-.010 (.002)***	-.123 (.013)***	-.0002 (.0003)
Educational Expenditures per capita	.00004 (1.00e-05)***	.00004 (1.00e-05)***	8.80e-06 (4.80e-06)*
Kindergarten Attendance	-.167 (.013)***	-.135 (.015)***	.006 (.011)
Sec. School gross enrollment ratio	-.116 (.014)***	-.083 (.016)***	-.026 (.008)***
Number of elementary school students	-7.20e-06 (1.10e-06)***	-.00006 (6.60e-06)***	1.00e-07 (2.00e-07)
Share of community schools	-.010 (.027)	.006 (.029)	.003 (.015)

Notes: Standard errors are reported in the parentheses. *** denotes significance at the 1% level, ** at the 5% level and * at the 10%. The treatment variable is at municipality level and it is defined as a share of students attending schools smaller than 71 students.

Table 4: The 2009 Reform and Changes in the School Network

Dependent Variable:	*School Handover [= 1]*		*School Closure [= 1]*	
	(1)	(2)	(3)	(4)
	Unit of Observation: Municipality-Year			
Panel A : Total				
The effect of treatment	.105	.096	.174	.179
	(.020)***	(.020)***	(.029)***	(.029)***
Observations	12385	12355	12385	12355
Number of municipalities	2477	2475	2477	2475
Overall R^2	.015	.00002	.028	.0002
Mean of dep. variable	.013	.013	.041	.041
Panel B : Rural subsample				
The effect of treatment	.109	.098	.176	.180
	(.021)***	(.020)***	(.030)***	(.030)***
Observations	11250	11221	11250	11221
Number of municipalities	2250	2248	2250	2248
Overall R^2	.017	.0003	.028	.015
Mean of dep. variable	014	.014	.042	.042
Panel C : Urban subsample				
The effect of treatment	.920	.923	.844	.746
	(.435)**	(.424)**	(.538)	(.538)
Observations	1135	1134	1135	1134
Number of municipalities	227	227	227	227
Overall R^2	.014	.0006	.028	.037
Mean of dep. variable	.01	.01	.028	.028
General Characteristics	No	Yes	No	Yes
Educational Characteristics	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes

Notes: Robust and clustered at the municipality level standard errors are reported in parentheses. *** denotes significance at 1% level, ** at 5% level and * at 10%. Table presents the coefficient of interaction between the treatment variable and the dummy *After* from regression (1). The general covariates include: unemployment rate, log of total municipality expenditure per capita, population, population density and a share of people aged 0-18. The educational covariates include: kindergarten attendance, secondary school gross enrollment ratio and log of municipality educational expenditures per capita. Dependent variables are the dummies, which take value one if in given year there was an episode of school handover (columns 1 and 2) or school closure (columns 3 and 4).

Table 5: Main Results

	Dependent Variable: standardized 6th grade exam			
	(1)	(2)	(3)	(4)
	Unit of Observation: School-Year			
Panel A : Total				
The effect of treatment	-.059	-.093	-.056	-.083
	(.064)	(.067)	(.064)	(.066)
Observations	64964	64544	64907	64488
Number of schools	9846	9846	9838	9838
Overall R^2	.001	.085	.003	.085
Mean of dep. variable	-.004	-.011	-.004	-.011
Panel B : Rural subsample				
The effect of treatment	-.051	-.044	-.037	-.029
	(.069)	(.070)	(.068)	(.068)
Observations	47535	47480	47478	47424
Number of schools	7144	7144	7136	7136
Overall R^2	.0005	.026	.0008	.026
Mean of dep. variable	-.15	-.15	-.15	-.15
Panel C : Urban subsample				
The effect of treatment	-2.282	-2.300	-2.282	-2.309
	(.678)***	(.687)***	(.675)***	(.683)***
Observations	17429	17064	17429	17064
Number of schools	2702	2702	2702	2702
Overall R^2	.002	.095	.01	.096
Mean of dep. variable	.395	.377	.395	.377
General Characteristics	No	Yes	No	Yes
Educational Characteristics	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes

Notes: Robust and clustered at the municipality level standard errors are reported in parentheses. *** denotes significance at 1% level, ** at 5% level and * at 10%. Table presents the coefficient of interaction between the treatment variable and the dummy *After* from regression (1). The general covariates include: unemployment rate, log of total municipality expenditure per capita, population, population density and a share of people aged 0-18. The educational covariates include: kindergarten attendance, secondary school gross enrollment ratio and log of municipality educational expenditures per capita. Dependent variable is the average of the standardized 6th grade exam score at the school level.

Table 6: Main Results - Full Regressions

	Dependent Variable: standardized 6th grade exam		
Sample:	Total	Rural	Urban
	Unit of Observation: School-Year		
The effect of treatment	-.083	-.029	-2.309
	(.066)	(.068)	(.683)***
Unemployment Rate	-.004	-.005	.007
	(.003)*	(.003)*	(.006)
Log Expenditures per capita	-.010	-.017	.022
	(.028)	(.031)	(.066)
Population	.003	.016	.001
	(.001)**	(.008)**	(.0009)
Population 0-18	-.008	-.014	-.022
	(.006)	(.007)*	(.013)*
Population Density	-.001	1.787	-.002
	(.0004)***	(.476)***	(.0006)***
Log Educational Expenditures per capita	.053	.052	.096
	(.033)	(.035)	(.083)
Kindergarten Attendance	.0003	.00006	-.0006
	(.0005)	(.0005)	(.001)
Sec. School gross enrollment ratio	-.001	-.0008	-.001
	(.0005)**	(.0006)	(.001)
Constant	-.214	-.296	-.260
	(.317)	(.325)	(.753)
Observations	64488	47424	17064
Number of schools	9838	7136	2702
Overall R^2	.085	.026	.096
General Characteristics	Yes	Yes	Yes
Educational Characteristics	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes

Notes: Robust and clustered at the municipality level standard errors are reported in parentheses. *** denotes significance at 1% level, ** at 5% level and * at 10%. Table presents coefficients from regression (1). Dependent variable is the average of the standardized 6th grade exam score at the school level.

Table 7: Heterogeneity: only Municipalities with Community Schools in 2008 (before the Reform)

	Dependent Variable: standardized 6th grade exam			
	(1)	(2)	(3)	(4)
	Unit of Observation: School-Year			
Panel A : Total				
The effect of treatment	-.132	-.220	-.238	-.282
	(.143)	(.151)	(.150)	(.158)*
Observations	17297	16928	17297	16928
Number of schools	2731	2731	2731	2731
Overall R^2	.004	.145	.092	.148
Mean of dep. variable	.294	.274	.294	.274
Panel B : Rural subsample				
The effect of treatment	-.169	-.132	-.198	-.159
	(.166)	(.168)	(.171)	(.173)
Observations	6445	6441	6445	6441
Number of schools	1004	1004	1004	1004
Overall R^2	.004	.019	.028	.02
Mean of dep. variable	-.129	-.128	-.129	-.128
Panel C : Urban subsample				
The effect of treatment	-2.898	-2.955	-2.907	-3.007
	(.979)***	(1.029)***	(.979)***	(1.028)***
Observations	10852	10487	10852	10487
Number of schools	1727	1727	1727	1727
Overall R^2	.004	.09	.006	.09
Mean of dep. variable	.545	.521	.545	.521
General Characteristics	No	Yes	No	Yes
Educational Characteristics	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes

Notes: Robust and clustered at the municipality level standard errors are reported in parentheses. *** denotes significance at 1% level, ** at 5% level and * at 10%. Table presents the coefficient of interaction between the treatment variable and the dummy *After* from regression (1). The general covariates include: unemployment rate, log of total municipality expenditure per capita, population, population density and share of people aged 0-18. Educational covariates include: kindergarten attendance, secondary school gross enrollment ratio and log of municipality educational expenditures per capita. Dependent variable is the average of the standardized 6th grade exam score at the school level.

Table 8: Heterogeneity: Schools Larger than 150/300 students

	Dependent Variable: standardized 6th grade exam			
	>150		>300	
	(1)	(2)	(3)	(4)
	Unit of Observation: School-Year			
Panel A : Total				
The effect of treatment	-.084	-.125	-.215	-.213
	(.083)	(.087)	(.145)	(.151)
Observations	34337	33975	16822	16606
Number of schools	4952	4952	2424	2424
Overall R^2	.004	.137	.007	.143
Mean of dep. variable	.133	.122	.287	.274
Panel B : Rural subsample				
The effect of treatment	-.039	-.032	-.070	-.018
	(.089)	(.090)	(.159)	(.163)
Observations	18950	18926	5928	5924
Number of schools	2725	2725	854	854
Overall R^2	.001	.037	.004	.029
Mean of dep. variable	-.103	-.103	-.012	-.012
Panel C : Urban subsample				
The effect of treatment	-2.165	-2.189	-1.941	-1.974
	(.669)***	(.669)***	(.703)***	(.712)***
Observations	15387	15049	10894	10682
Number of schools	2227	2227	1570	1570
Overall R^2	.004	.124	.005	.126
Mean of dep. variable	.424	.405	.449	.432
General Characteristics	No	Yes	No	Yes
Educational Characteristics	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes

Notes: Robust and clustered at the municipality level standard errors are reported in parentheses. *** denotes significance at 1% level, ** at 5% level and * at 10%. Table presents the coefficient of interaction between the treatment variable and the dummy *After* from regression (1). The general covariates include: unemployment rate, log of total municipality expenditure per capita, population, population density and share of people aged 0-18. Educational covariates include: kindergarten attendance, secondary school gross enrollment ratio and log of municipality educational expenditures per capita. Dependent variable is the average of the standardized 6th grade exam score at the school level.

Table 9: Robustness: The Municipality Average of the 9th Grade Exam Score (Placebo Experiment)

	Dependent Variable: standardized 9th grade exam			
	(1)	(2)	(3)	(4)
	Unit of Observation: Municipality-Year			
Panel A : Total				
The effect of treatment	.199	.211	.199	.211
	(.114)*	(.116)*	(.114)*	(.116)*
Observations	17294	17272	17294	17272
Number of schools	2472	2472	2472	2472
Overall R^2	.002	.015	.002	.015
Mean of dep. variable	-.0002	-.0006	-.0002	-.0006
Panel B : Rural subsample				
The effect of treatment	.206	.227	.206	.227
	(.119)*	(.120)*	(.119)*	(.120)*
Observations	15593	15573	15593	15573
Number of schools	2229	2229	2229	2229
Overall R^2	.0002	.008	.0002	.008
Mean of dep. variable	-.061	-.061	-.061	-.061
Panel C : Urban subsample				
The effect of treatment	.275	.521	.275	.521
	(1.058)	(1.004)	(1.058)	(1.004)
Observations	1694	1692	1694	1692
Number of schools	242	242	242	242
Overall R^2	.009	.076	.009	.076
Mean of dep. variable	.558	.555	.558	.555
General Characteristics	No	Yes	No	Yes
Educational Characteristics	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes

Notes: Robust and clustered at the municipality level standard errors are reported in parentheses. *** denotes significance at 1% level, ** at 5% level and * at 10%. Table presents the coefficient of interaction between the treatment variable and the dummy *After* from regression (1). The general covariates include: unemployment rate, log of total municipality expenditure per capita, population, population density and share of people aged 0-18. Educational covariates include: kindergarten attendance, secondary school gross enrollment ratio and log of municipality educational expenditures per capita. Dependent variable is the average of the standardized 6th grade exam score at the municipality level.

Table 10: Sorting: The Municipality Average of the 6th Grade Exam Score

	Dependent Variable: standardized 6th grade exam			
	(1)	(2)	(3)	(4)
	Unit of Observation: Municipality-Year			
Panel A : Total				
The effect of treatment	-.067	-.072	-.067	-.072
	(.100)	(.101)	(.100)	(.101)
Observations	17337	17301	17337	17301
Number of schools	2477	2475	2477	2475
Overall R^2	.004	.043	.004	.043
Mean of dep. variable	-.0002	-.0005	-.0002	-.0005
Panel B : Rural subsample				
The effect of treatment	-.044	-.017	-.044	-.017
	(.105)	(.105)	(.105)	(.105)
Observations	15748	15714	15748	15714
Number of schools	2250	2248	2250	2248
Overall R^2	.0003	.068	.0003	.068
Mean of dep. variable	-.083	-.083	-.083	-.083
Panel C : Urban subsample				
The effect of treatmen	-1.883	-1.831	-1.883	-1.831
	(1.154)*	(1.155)	(1.154)*	(1.155)
Observations	1589	1587	1589	1587
Number of schools	227	227	227	227
Overall R^2	.013	.093	.013	.093
Mean of dep. variable	.817	.815	.817	.815
General Characteristics	No	Yes	No	Yes
Educational Characteristics	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes

Notes: Robust and clustered at the municipality level standard errors are reported in parentheses. *** denotes significance at 1% level, ** at 5% level and * at 10%. Table presents the coefficient of interaction between the treatment variable and the dummy *After* from regression (1). The general covariates include: unemployment rate, log of total municipality expenditure per capita, population, population density and share of people aged 0-18. Educational covariates include: kindergarten attendance, secondary school gross enrolment ratio and log of municipality educational expenditures per capita. Dependent variable is the average of the standardized 6th grade exam score at the municipality level.

10.14754/CEU.2016.03

Table 11: School Spending

Dep. Variable:	Share of Renovation Costs		Share of Teacher Salaries		Green Areas per Student		Sport Area per Student	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Unit of Observation: School-Year				
Panel A : Total								
The effect of treatment	9.274	11.193	.012	.006	1.678	1.764	.002	-.092
	(21.178)	(20.575)	(.011)	(.011)	(2.124)	(2.247)	(1.049)	(1.095)
Observations	45639	45371	45454	45187	45596	45331	45596	45331
Number of schools	9191	9183	9193	9185	9197	9189	9197	9189
Overall R^2	.002	.0006	.002	.002	.002	.006	.00006	.009
Mean of dep. variable	75.84	75.886	.648	.648	22.945	22.991	13.484	13.504
Panel B : Rural subsample								
The effect of treatment	15.106	16.501	-.002	-.004	1.091	1.300	-.875	-.778
	(22.452)	(22.684)	(.011)	(.011)	(2.275)	(2.336)	(1.126)	(1.148)
Observations	33393	33300	33265	33172	33380	33289	33380	33289
Number of schools	6690	6682	6692	6684	6699	6691	6699	6691
Overall R^2	.002	.00009	.001	.0006	.0002	.004	.00003	.004
Mean of dep. variable	77.308	77.448	.646	.646	27.181	27.188	15.869	15.869
Panel C : Urban subsample								
The effect of treatment	56.210	4.940	.072	.055	-7.156	-11.123	.739	.826
	(214.801)	(215.383)	(.128)	(.120)	(12.049)	(11.823)	(16.186)	(15.747)
Observations	12246	12071	12189	12015	12216	12042	12216	12042
Number of schools	2501	2501	2501	2501	2498	2498	2498	2498
Overall R^2	.003	.004	.004	.022	.00008	.0004	.0006	3.88e-07
Mean of dep. variable	71.835	71.578	.653	.654	11.371	11.388	6.969	6.965
General Characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Educational Characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust and clustered at the municipality level standard errors are reported in parentheses. *** denotes significance at 1% level, ** at 5% level and * at 10%. Table presents the coefficient of interaction between the treatment variable and the dummy *After* from regression (1). The general covariates include: unemployment rate, log of total municipality expenditure per capita, population, population density and share of people aged 0-18. Educational covariates include: kindergarten attendance, secondary school gross enrolment ratio and log of municipality educational expenditures per capita.

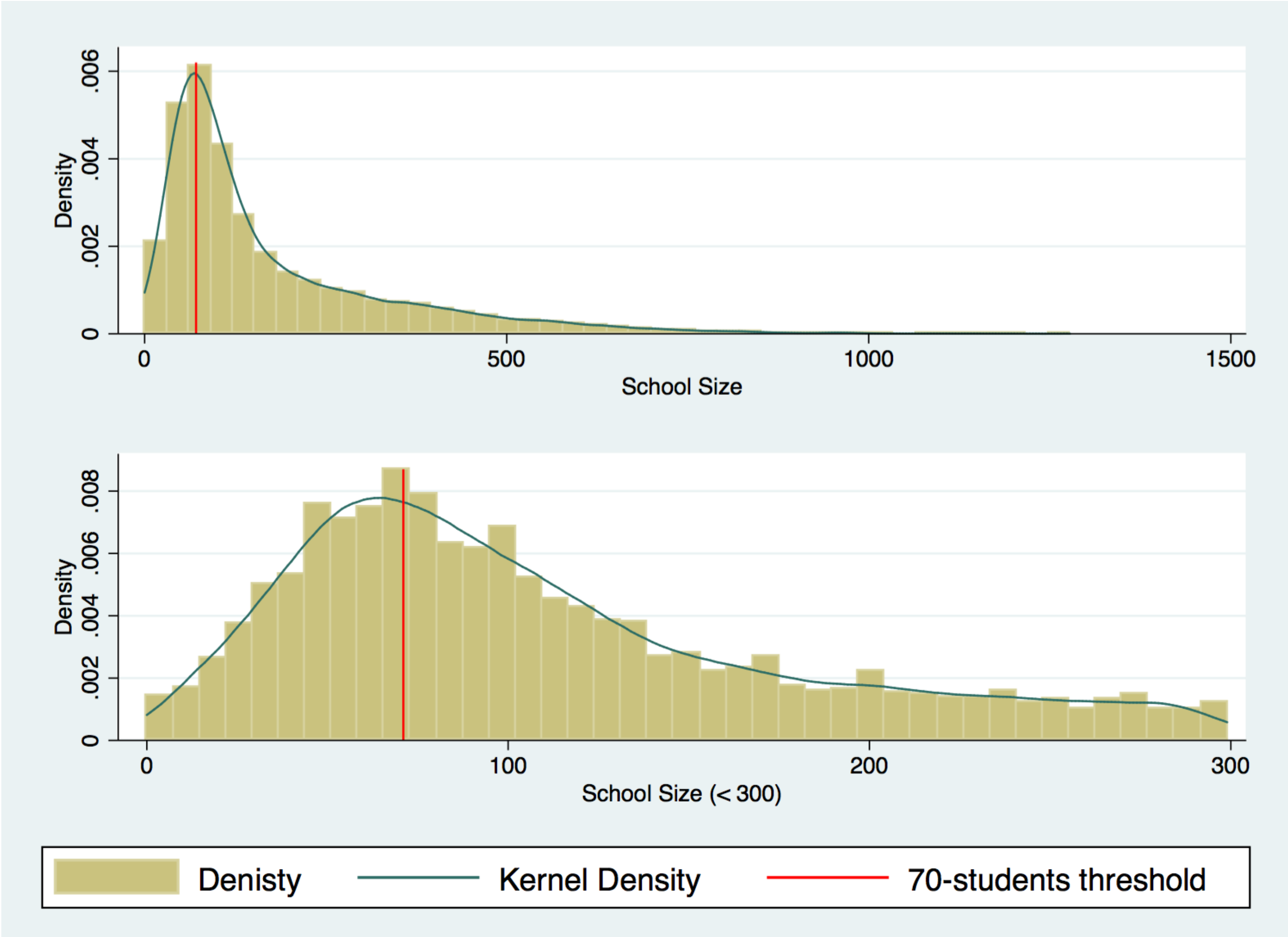


Figure 1: Distribution of School Size in 2008

Note: the plot presents the distribution of school size in 2008. The upper panel is for all schools, the bottom panel only for schools smaller than 300 students.

10.14754/CEU.2016.03

10.14754/CEU.2016.03

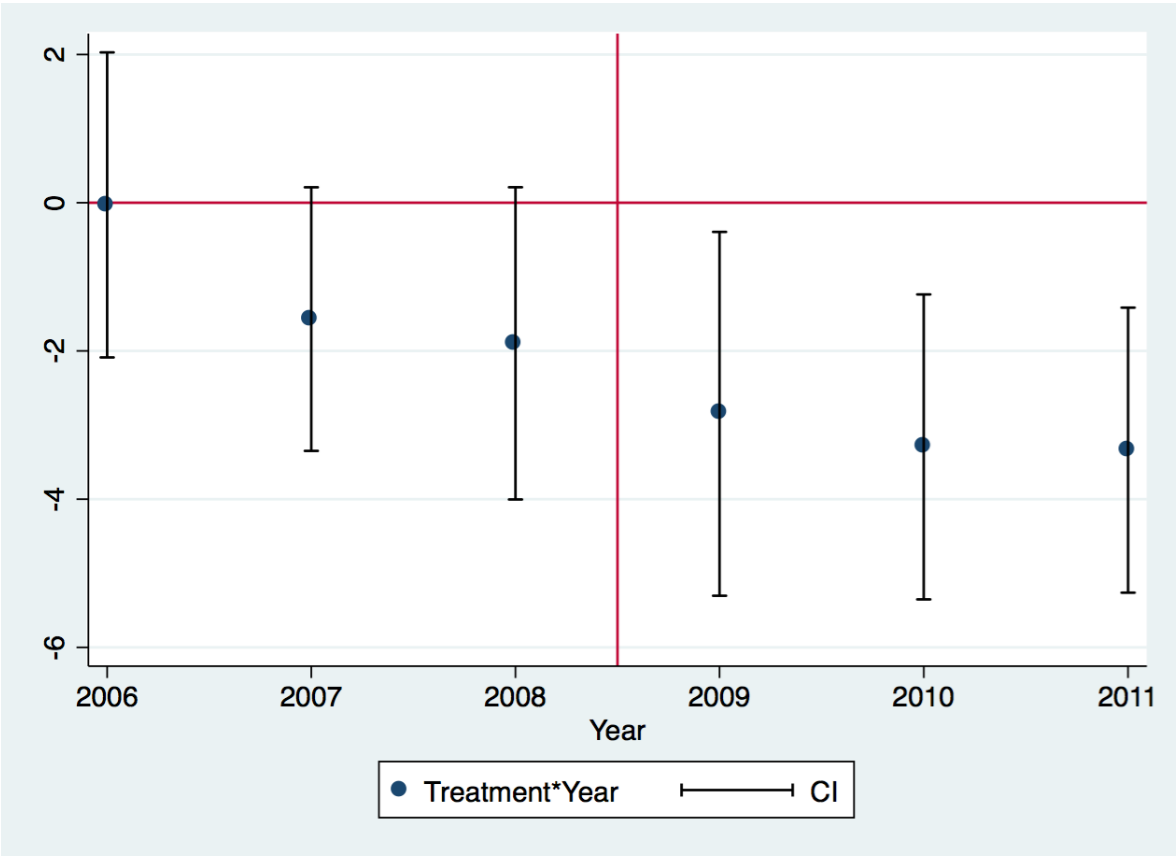


Figure 2: The Generalized Model

Note: the plot presents the estimates and 95% confidence intervals of interaction between the treatment variable and year fixed effects. The coefficient were obtained from a regression of the school-level 6th grade standardised exam scores on the set of interactions and time fixed effects. The confidence intervals calculated using clustered and robust standard errors.

10.14754/CEU.2016.03

Chapter 3
School Competition and Sorting of Students Within a School*

Existing literature shows that school competition might lead to sorting of students between schools. This is a first study to estimate the effect of school competition on sorting within a school (across classes). The identification strategy is based on a two-stage design of the Polish Comprehensive Education, which allows to isolate an exogenous change in student mobility. In addition, I use a novel measure of student socio-economic characteristics - Raven's Progressive Matrix test score. The results show that school competition leads to a higher sorting of students within a school and between schools. I investigate two explanation of the effect on sorting within a school: the demand for peer quality (Epple et al., 2002) and the demand for teachers (Clotfelter et al., 2005). The data point to the importance of the former mechanism, i.e. the demand for high quality peers that motivates school principals to create high tracks within a school.

*I thank Sascha O. Becker, David Card, Roman Dolata, Bernhard Enzi, Hedvig Horvát, Gábor Kézdi, Randy Filer, Sergey Lychagin, Jesse Rothstein, Christopher Walters, the participants of seminars at Central European University and University of California at Berkeley for their comments and suggestions. I acknowledge the hospitality of the Department of Economics, CLE, IRLE at UC Berkeley. I gratefully acknowledge financial support from the CERGE-EI/GDN Grant RRC16+12. All errors are mine.

1 Introduction

It has been argued that school competition might motivate school principals to improve school quality (Friedman, 1955; Hoxby, 2000). Based on this premise many policies, such as school vouchers or school autonomy, have been recently proposed to accelerate student mobility (Gibbons, Machin and Silva, 2008).[1] On the other hand, a number of studies shows that school competition and student mobility might lead to sorting of students between schools (Epple and Romano, 1998; Ladd and Fiske, 2000; Hsieh and Urquiola, 2006; Nechyba, 2006; Böhlmark, Holmlund, Lindahl et al., 2015). However, we know less about the effect of school competition on sorting within a school (across classes). This is surprising given the importance of class assignment (Kalogrides, Loeb and Béteille, 2013; Collins and Gan, 2013) and that student sorting is not neutral for the performance of students and might violate educational equality of opportunity (Meghir and Palme, 2005; Kremer, Duflo and Dupas, 2011; Figlio and Page, 2002).

In this paper I investigate the effect of school competition on sorting within a school and between schools. In order to isolate an exogenous change in student mobility, which increases school competition, I exploit a two-stage design of the Polish Comprehensive education. For measuring sorting I use the fraction of the variance of Raven's Progressive Matrix test score explained by school or class levels. Raven's score is a measure of general intelligence, which is determined by student genetic abilities and socio-economic background. It is fixed since early childhood, which ensures that the only source of class/school homogeneity is sorting of students. The results show that school competition leads to more homogeneous classes and schools.

Next, I focus on the potential mechanisms explaining the effect of school competition on sorting within a school. A theoretical model by Epple, Newlon and Romano (2002) predicts that high track might be used to attract high-skill or high-income students (the demand for peer quality). In a complementary work, Clotfelter, Ladd and Vigdor (2005) claim that it might be also used to attract high-skilled teachers (the demand for teachers). Using data on school characteristics I empirically test these two channels. The results point out to the importance of the demand for peer quality channel.

This paper might be useful for policymakers who wish to use school competition as a mean to improve quality of schools, but also wish to avoid its negative distributional consequences. The results underlines the importance of school principals' incen-

[1] In the USA school choice is promoted by expansion of the Charter School laws (Kern, Thukral and Ziebarth, 2012). In Poland there has been a movement for fostering creation of parent-lead autonomous Community Schools (Sochocki, 2011).

tive structure. Classroom assignment, by creating classes with high *level* of peer quality, might be used by principals to attract high-achievers or high-income students. This could be weaken by the incorporation of value added estimates of school performance into principals' objectives, as it motivates them to compete also for low-background or low-performing students (MacLeod and Urquiola, 2009). Even though the value-added based accountability has been heavily discussed, not much attention has been paid to the potential distributional effects (Rothstein, 2009; Angrist, Pathak and Walters, 2011; Chetty, Friedman and Rockoff, 2014). Alternative policy could be to link school vouchers with the socioeconomic background, for instance to offer them only to students with low income [2] On the other hand, abolishing the teacher collective bargaining agreements allows school principals to compete based on wages rather than composition of students. Nevertheless, in this study I do not find strong evidence for the demand for teachers mechanism.

The paper is organized as follows. In the second section I discuss the organization of the Polish education system. The third section is devoted to the identification strategy. The fourth section presents empirical methodology and data. In the fifth section I show the main results and robustness checks. In the sixth section I discuss in more detail and empirically investigate the effect of school competition on sorting within a school. Finally, in the sixth section I conclude.

2 Institutional Background

The Polish comprehensive education is compulsory and consists of six years of elementary school (ISCED 1), which is followed by three years of lower secondary school *gimnazjum* (ISCED 2). Elementary school and *gimnazjum* usually serve the same community of students, but they are separated entities, with different managerial and teaching bodies. After finishing the comprehensive part, student may finish their education or continue in academic, mixed or vocational higher secondary school (ICED 3).

The admission to elementary school and *gimnazjum* is identical. It is based on catchment areas, which means that every student has a right to attend an assigned local public school and this school has to accept her. Because there are more elementary schools than *gimnazja*,[3] the catchment area for the later is usually larger and contains

[2]This policy is in effect in Chile and the Netherlands, see Böhlmark et al. (2015).

[3]Most of elementary schools were constructed during the past 50 years, while *gimnazja* only after 1999. The network of elementary schools thus reflects the past demographic situation and it is considered as too dense. The network of *gimnazja*, in turn, is more "rational" in the sense that it is better

the catchment areas from several elementary schools. Table 1 shows ratio of elementary schools to *gimnazja* in a rural-urban breakdown. In the rural areas there are on average 2.3 elementary school per *gimnazjum* and almost 1.5 in the urban areas. As an alternative to the local school, parents may request a place in an under-subscribed non-local school, but the admission is not granted. The are no universal recruitment rules for non-local students. Each school's policy is determined by school principal and a recruitment committee, which usually consists of selected teachers and school psychologist.

In theory the classroom assignment is similar for elementary schools and *gimnazja*. There are no universal rules and the assignment is determined by school principals and the recruitment committee, which create own list of criteria (often conflicting). As for *gimnazja*, the most common rules are that classes should be equal in terms of student performance, gather students with similar GPA, with similar foreign language proficiency or from the same neighborhood (Szmigel, 2013). Usually parents have a right to suggest an alternative class assignment. Elementary schools, in turn, cannot sort students based on their performance (it is unknown), so usually use place of living, gender, special needs, date of birth and parental preferences as criteria. In 2010 there were no limits for classroom size.[4] Importantly, the assignment is fixed across grades, subjects and reallocation across classes is allowed only in special situations. Hence, the peer composition of classes is usually constant within each stage of education.

During the comprehensive education, students are examined by two standardized, externally graded and obligatory examinations: a low stake after elementary school (6th grade) and a high stake after *gimnazjum* (9th grade).[5] The school averages from these exams are published in various unofficial school rankings. The only official measure of school quality is the school-level educational value added, but it is only available for *gimnazja*.[6]

There are clear economies of scale for school principals of elementary school and *gimnazja*. All Polish public schools are financed by the central government through a subsidy. In theory this amount should be sufficient to cover all expenditures on education, excluding investments and pre-school education. In practice, however, it covers

adjusted to the current demographic needs. In addition, elementary schools serve younger children for whom distance to a school matter more than for older children.

[4]Since 2015 the rules in elementary schools have been unified and are based on the date of birth with an option for parents to request an alternative assignment. Since 2013 a class in grades I-III can have maximum 25 students.

[5]The 9th grade exam serves as a basis for the admission into the higher secondary education (ISCED 3).

[6]It is widely available only since 2009. The Ministry of Education publishes also annual level-based rankings of the higher secondary school.

only around 50-70% of the costs (Herbst, Herczyński and Levitas, 2009; Instytut Badań Edukacyjnych, 2011) and the rest is covered by local governments. Since the governmental subsidy is stuck to the student (the money goes with her), the amount of school funds depends on enrollment. In addition, school principal might gain more local political power from larger schools, which might be crucial for securing additional funds from the municipality government. These, together with a negative population growth threatening existence of many public schools, motivate schools to increase their enrollment. Classroom assignment is a potential mean of competition for students. Sorting can be used to attract more high achievers, to attract high quality teachers or to improve quality of education by tailoring the teaching practice. On the other hand, mixing students might improve educational equality of opportunity and thus be preferred by principals and policymakers.

The teacher wages and general employment conditions are mostly determined by municipality governments (not by school principals) in compliance with the universal collective bargaining agreements (*Karta Nauczyciela*). Salary has to be at least as large as a minimum wage determined for each teacher's rank in *Karta Nauczyciela*.[7] In addition, teachers may receive extra salary for working over-time, monetary awards and other non-monetary benefits, for instance accommodation in school's social apartments.

[7]In 2015 the minimum monthly gross wages ranged from 1513 PLN (340 EUR) to 3109 (700 EUR). Additionally, local municipalities have to make sure that the average *total* gross salary for each tacher's rank within municipality is at least as large as specified in *Karta Nauczciela*. In 2015 these averages ranged from 2717 PLN (612 EUR) to 5000 PLN (1126 EUR).

Table 1: Comparison of the Rural and Urban areas in Poland.

Variable	Rural	Urban
Total Numbers		
Elementary Schools (2012)	8604	4092
Gimnazja (2012)	3722	2748
Elementary School per *Gimnazja* (2012)	2.31	1.49
Averages for Municipalities		
Elementary School per km^2 (2010)	0.028	0.21
Gimnazjum per km^2 (2010)	0.01	0.18
Children per Elementary School (2010)	156	337
Children per *Gimnazjum* (2010)	186	207
Public Transportation per km^2 (2007)	.071	1.656
Tertiary Education Share (2002)	4%	11%
Population Density (2010)	166	1676
Population (2010)	10067	156004
Number of Municipalities	2386	93

Source: the Central Statistical Office of Poland and Herczyński & Sobotka (2013). Urban = population > 50 000.

3 Identification Strategy

The effect of school competition on student sorting is likely to be confounded with the effects of other parallel social processes. Because of neighborhood quality, local economic conditions or historical accidents, similar people tend to live together (Tiebout, 1956). At the same time, housing prices might be influenced by local school characteristics, which further reinforces self selection (Figlio and Lucas, 2004; Kane, Riegg and Staiger, 2006). Consequently, a regression of a measure of school heterogeneity on a measure of competition will be biased by residential sorting if school admission process is based on catchment areas and places with higher school competition have also higher residential sorting. In order to solve this problem, I exploit the two-stage design of the Polish comprehensive education, which under the set of three assumption, allows me to isolate an exogenous change in school choice. The change in school choice, in turn, translates into higher school competition.

The idea is to compare sorting of students across classes/schools at the entrance to *gimnazjum* with sorting at the entrance to elementary school. This cancels out the influence of residential sorting and other grade-invariant local characteristics. The remaining difference is due to the higher student mobility and general across-stage changes in classroom assignment or mixing effect of larger catchment areas. However, in the areas with

10.14754/CEU.2016.03

Higher Student Mobility

	1st grade: Elementary School	7th grade: Gimnazjum
Competition (low cost)	Residential Sorting	Residential Sorting + Classroom Assignment + **Ability Sorting across Classrooms**
No Competition (high cost)	Residential' Sorting	Residential' Sorting + Classroom Assignment

Figure 1: The Identification Strategy - Sorting Within a School

high costs of school choice the difference will be to a lesser extent affected by student mobility. One can thus isolate the effect of school competition (higher student mobility) on sorting by comparing how sorting differs across places with different costs of school choice. The identification strategy is summarized in Figures 1 and 2. Each cell lists the forces driving the classroom (Figures 1) and school (Figure 2) homogeneity across stages of education and locations with high and low potential for school competition. This design can be seen as an example of the difference-in-difference methodology. "Treatment" is an increase in school competition induced by the increase in student mobility, "treatment group" is an area with a high potential for school competition (low school choice costs, the first rows), and "before and after" are the first and second stages of the Polish comprehensive education respectively (the columns). This strategy produces a causal effect of school competition on student sorting when the three assumptions hold.

Assumption 1. *Students who enter gimnazjum (the seventh grade) are more likely to use school choice than students who enter elementary school (the first grade).*

This first assumption is motivated by three observations. Firstly, students who enter *gimnazjum* are older so it is easier for them to travel larger distances to an alternative school. Secondly, students are tested at the end of elementary school and their performance is a common knowledge. Contrary to this, skills of students entering elementary school are unknown. Smaller informational constraints might motivate students to select a non-local school and allow school principals to effectively select applicants. Thirdly, a catchment area of *gimnazjum* usually contains catchment areas of local elementary schools. Consequently, students entering *gimnazjum* are facing larger catchment areas

10.14754/CEU.2016.03

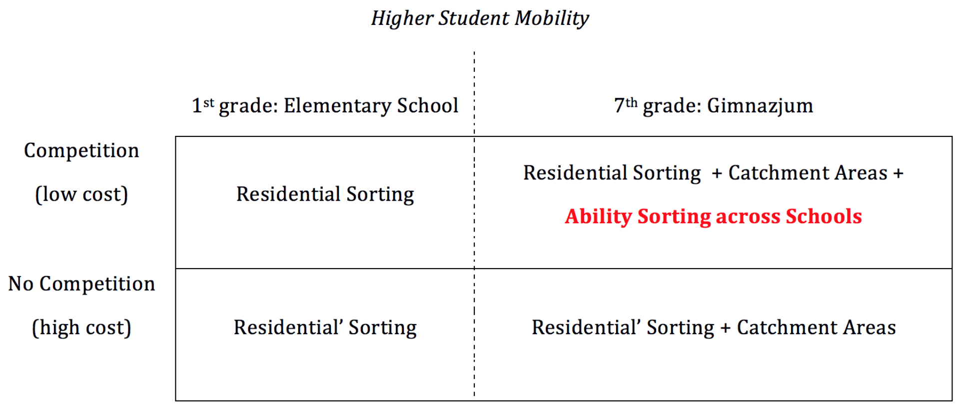


Figure 2: The Identification Strategy - Sorting Between Schools

and school composition of their local *gimnazjum* will to a lesser extent (than in the case of elementary schools) reflect residential composition of their neighborhood. Lesser similarity might provide smaller incentives to attend a local school.

This assumption might be not valid when elementary school students have more schools within their reach, making competition more likely. Indeed, as reported in Table 1, there are more elementary schools than *gimnazja*. This is especially visible in rural areas, which are areas with a low potential for school competition. The second assumption is needed in order to obtain a control group.

Assumption 2. *The difference in student mobility across educational stages is irrelevant in areas with a low potential for school competition (i.e. rural areas).*

The motivation for this assumption is that in rural areas school variety is limited and cost of sending a child to a non-local school is significant (Dolata, 2008). The cost of school choice includes a transportation cost, missing links with peers from the neighborhood or limited possibilities of grass-root actions with other parents. Table 1 shows that indeed rural municipalities, i.e. those with population lower than 50 000, have three (ten) times sparser network of elementary schools (*gimnazja*), twenty-three times sparser network of public transportation and ten times smaller population density.

The third identifying assumption is different for the analysis of sorting within a schools and between schools.

Assumption 3.a. *(for sorting within a school) In the absence of an increase in school competition, a change in class assignment between elementary schools and gimnazja is* *the same in areas with different potential for school competition (i.e. urban and rural areas).*

Schools might have various reasons to sort or mix students, which are unrelated to school competition. This assumption says that these should be similar in areas with different potential for competition. This is supported by qualitative evidence (discussed in Section 6.2), as generally principals want to balance the class composition in terms of performance, but at the same time keep students from the same area together.

Assumption 3.b. *(for sorting between schools) The change in the size of catchment areas between elementary schools and gimnazja leads to the same level of between-school student mixing in urban and rural areas.*

Unlike assumption 3.a, this assumption is not likely to be satisfied. Student mixing should be more intensive in the rural areas as there are more elementary schools per *gimnazjum* than in the urban areas (see Table 1). I try to account for this difference in the result section.

4 Estimation and Data

The first part of this section proposes an estimator for a change in sorting of students across stages of education, which is robust for differences in the size of catchment areas. I develop separate methods for capturing changes in sorting between and within schools. In the second part of this section I discuss the data and the main measure of socio-economic background.

4.1 Estimation

I focus firstly on sorting between schools. Consider some measure of socio-economic background (SEB): y_{ics} of student i from class c and school s. It can be decomposed into the population mean μ, the school-level deviation from that mean u_s, the class-level deviation from the school mean u_c and the residual component e_{ics}.

$$y_{ics} = \mu + u_c + u_s + e_{ics} \tag{1}$$

By construction, the variance of the SEB variable at stage t (either Gimnazjum - *gim* or Elementary School -*es*) is a sum of the variance of the school-level component, the variance of the class-level component and the residual variance.

$$Var_t = Var_{s,t} + Var_{c,t} + Var_{e,t} \tag{2}$$

For a given educational stage, an intensity of sorting between schools can be defined as a ratio of the school-level variance to the total variance: $\frac{Var_{s,t}}{Var_t}$. The change across educational stages is:

$$\Delta Var_s = \frac{Var_{s,gim}}{Var_{gim}} - \frac{Var_{s,es}}{Var_{es}} \tag{3}$$

A change in sorting within a school can be captured in a similar way, except that one has to correct for the differences in catchment areas between elementary school and *gimnazjum*. Generally, an intensity of sorting within a school is defined as a ratio of the class-level variance to the total variance: $\frac{Var_{c,t}}{Var_t}$. Ignoring the catchment area problem, the change between educational stages is simply: $\frac{Var_{c,gim}}{Var_{gim}} - \frac{Var_{c,es}}{Var_{es}}$

The problem arises because the catchment areas are larger for *gimnazja* than for elementary schools. When there are no changes in class composition at the transition between stages, the fraction of variance explained by the school-level drops and the fraction explained by the class-level might increase correspondingly. To see this, suppose that there is just one class per elementary school and students have exactly the same classmates during both elementary school and *gimnazjum*. Because of the nested catchment areas, students from several elementary schools will go to one *gimnazjum*, and each class in that *gimnazjum* will consist of students coming from the same elementary school. This implies that the relative importance of the class-level ($\frac{Var_{c,t}}{Var_t}$) increases, even though there was no change in student sorting across classrooms.[8] To correct for this problem one can adjust for the negative change in the fraction of the variance explained by the school-level. I propose a following measure of the change in sorting within a school:

$$\Delta Var_c = \frac{Var_{c,gim}}{Var_{gim}} - \frac{Var_{c,es}}{Var_{es}} + \mathbb{1}_{[\Delta Var_s < 0]} \Delta Var_s \tag{4}$$

where $\mathbb{1}_{[a]}$ is an indicator function, which takes value zero if expression a is not true and one when it is true, that is a change in the fraction of variance explained by the school-level is negative. Intuitively, the aforementioned problem arises only when *gim-*

[8]The other way of looking at this problem is to realize that, in this scenario, schools at the elementary school stage become classes at the *gimnazjum* stage. With one class per elementary school there is no difference between labels: "school" and "class". Although there is no change in the class composition at the transition to *gimnazjum*, the distinction between "school" and "class" starts to matter. This is because groups of students, which were "classes/schools" at the elementary stage, becomes "classes" at the secondary stage.

nazja have larger catchment areas than elementary schools and their ratio $\frac{Var_{s,t}}{Var_t}$ is lower. When there is no change in class composition, but catchment areas are larger for secondary school, $\frac{Var_{c,gim}}{Var_{gim}} - \frac{Var_{c,es}}{Var_{es}} = -\Delta Var_s$ and thus ΔVar_s should be subtracted in order to obtain value of zero. If the catchment areas are the same or sorting across schools overbalances their effect, a simple difference between the fraction of the variance explained by the class-level captures the effect of interest.

To isolate the effect of school competition, I compare changes in sorting in areas with different potential for school competition (different cost of exerting school choice). I assume that in rural areas the costs of school choice are so high that everybody follow their local school, whereas in urban areas students can go to a non-local one. The effect of school competition on sorting within and between schools can be defined as:

$$\Delta Var_c^{URBAN} - \Delta Var_c^{RURAL} \tag{5}$$

$$\Delta Var_s^{URBAN} - \Delta Var_s^{RURAL} \tag{6}$$

I use a multilevel mixed-effects linear regression framework (also called a hierarchical linear model) to estimate the proportion of variance of the SEB variable explained by the class and the school levels.

4.2 Data

My main measure of the SEB is Raven's Progressive Matrix test. It is designed to capture two abilities: "(a) eductive ability [...] - the ability to make meaning out of confusion, the ability to generate high-level, usually nonverbal, schemata which make it easy to handle complexity; and (b) reproductive ability - the ability to absorb, recall, and reproduce information that has been made explicit and communicated from one person to another" (Raven, 2000, p.2). In other words, eductive and repructive abilities allow to understand concepts and to learn new material. They are components of an underlying general mental ability, which is also called (after its creator) the Spearman g factor (Jensen, 1998). The test usually consists of 4x4 3x3 or 2x2 matrix of figures at each entries except the lowest diagonal which is empty. Figures in each row are following the same pattern and the task of the subject is to identify the missing element according to this pattern. Importantly, Raven's test score is determined only by genetic, parental and environmental conditions during early childhood (Brouwers, Van de Vijver and Van Hemert, 2009). Any post-kindergarten determinants of education, such as peer effects, school inputs, teacher quality or parental investments should not matter. Consequently, the only reason why students might have similar level of Raven's score is self-selection. The advantage of Raven's score is that it includes characteristics affecting sorting of students, such as genotype, which are not easily captured by other commonly used measures (e.g., mother's education).

The data are drawn from the sample of Polish students collected by the Educational Value Added Team.[9] The cross-section is from 2010 and consists of 5600 first-graders and 5567 seventh-graders (which is an entry grade of *gimnazjum*) from 330 randomly drawn public schools in Poland.[10] The main outcome variable and measure of background characteristics is a standardized (separately for the first and seventh graders) cumulative score from Raven's Progressive Matrix test. For each student i from grade g, I calculate Raven's z-score, that is:

$$zscore_{ig} = \frac{score_{ig} - \overline{score_g}}{sd(score_g)}$$

where $score_{ig}$ is raw Raven's score and $sd(score_{ig})$ is a standard deviation of Raven's score for each grade. In addition, a set of student, parental, teacher, school and municipality - level characteristics is available. Importantly, it includes questions about each school's sorting practices. All the statistics used in the paper are weighted using an appropriate weighting scheme, thus the results should be interpreted as representative for the corresponding Polish populations. Table 2 summarizes the available sample.

A potential test for the claim that Raven's score is not affected by education is to regress mother's and father's education on Raven's score, a dummy denoting observations from *gimnazjum* (the seventh grade) and an interaction between the two. If Raven's score is not affected by education there should be no difference in correlation between parental education and Raven's score for the first and seventh graders. Table 3 Columns (1) and (2) show that while there is a positive correlation between mother's and father's education and Raven's score, it is not significantly different between the first and seventh grades.

[9]The Project was funded by the European Union under the European Social Found and was ran by the Central Examination Board until September 2012. Since October 2012 the project is run by Educational Research Institute in Warsaw.

[10]For elementary schools, the population were public schools with first grades larger than 10 students. For *gimnazja*, the population were public schools with seventh grades larger than 20 students.

10.14754/CEU.2016.03

Table 2: Descriptive Statistics of The Sample

	Elementary School					*Gimnazjum*				
Variable	Obs.	Mean	St. Dev.	Min	Max	Obs.	Mean	St. Dev.	Min	Max
Full sample										
Raw Raven's score	5589	27.42	8.38	1	59	4907	45.27	7.58	9	60
Respondents per school	5749	36.17	10.04	8	56	4916	34.39	7.14	10	58
Respondents per class	5749	19.35	4.17	8	30	4916	17.81	4.09	6	30
Numbef of schools	180					150				
Urban sample										
Raw Raven's score	2103	29.16	8.31	9	55	1524	46.32	7.48	9	60
Respondents per school	2181	39.83	8.25	10	56	1526	35.48	8.53	10	58
Respondents per class	2181	20.2	4.23	8	28	1526	18.26	4.71	8	30
Numbef of schools	58					46				
Rural sample										
Raw Raven's score	3486	26.38	8.24	1	59	3383	44.79	7.57	10	60
Respondents per school	3568	33.94	10.37	8	50	3390	33.9	6.36	15	49
Respondents per class	3568	18.84	4.06	8	30	3390	17.6	3.75	6	28
Numbef of schools	122					104				

Notes: Note: The descriptive statistics are calculated for the sample, not for the population, therefore no weighting is used.

On the other hand, Column (3) shows that there is a more positive correlation between Raven's and desired education for a child at the seventh grade than at the first grade. This might be explained by lesser informational constrains faced by parents at the entrance to *gimnazjum*. Since, as reported in Column (4), there is a positive correlation between the sixth grade GPA and Raven's score, students with higher Raven's score are on average better performers and their parents might desire a higher level of education for them. Student performance is unknown for parents at the entrance to elementary school and thus the correlation between Raven's score and the desired education is significantly lower.

Table 3: Correlations between Raven's score and various outcomes.

Dependent Variable:	Mother's Education	Father's Education	Desired Education for a Child	6th grade GPA
	(1)	(2)	(3)	(4)
Raven's Score	.557	.543	.464	.532
	(.042)***	(.040)***	(.035)***	(.017)***
Gimnazjum	-.265	-.219	-.352	
	(.072)***	(.072)***	(.065)***	
Raven's Score *X Gimnazjum*	-.019	-.008	.370	
	(.051)	(.050)	(.042)***	
N	10320	10167	10376	4896
Estimator	OLogit	OLogit	OLogit	OLS

Notes: The table shows regressions of the depended variables on the standardized Raven's Progressive Matrix Test score, a dummy indicating observation from the seventh grade - *Gimnazjum* (excluded category is the first grade - elementary school), and the interaction between them. Mother's and Father's Education are categorical variables, which take valuess between 1 and 9, where 1 is unfinished elementary education and 9 is PhD. Desired Education for a Child is a categorical variable, which takes values between 1 and 7, where 1 is vocational education and 7 PhD. 6th grade GPA is the average of grades from various subjects, it ranges between 2 and 6, where 2 is the worst. Robust and corrected for the survey design standard errors are reported in the parentheses. In columns (1) to (3) the numbers show the coefficients from the Ordered Logit regression. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

5 Variance Decomposition of Raven's Score

The first part of this section presents the variance decomposition of the standardized Raven's score. Next, using methodology described in the previous section, I translate the results into the effect of school competition on sorting of students. The second part shows the robustness checks.

5.1 Results

Table 4 presents the proportion of variance of the standardized Raven's score explained by the school and class levels, along with residual variation, in breakdown by elementary school *gimnazjum*, and by urban and rural areas. The proportions and standard errors are estimated using a mixed effect model, the survey weights are taken into account. Figures 3 and 4 visualize the results.

Firstly I focus on urban areas. At the entrance to elementary school, the school and class levels explain 13% and 1% respectively of the Raven's score variation. At the entrance to *gimnazjum*, both proportions increase to 28% and 9% respectively. These results show that *gimnazja* and *gimnazja*'s classes are more homogeneous than in the case of elementary school. Consequently, the unexplained (residual) proportion of variance drops from 86% to 63%. Using Assumption 1, I argue that the increase in homogeneity is due to student's increased willingness to exert a school choice (higher school competition). Because of the economies of scale, school principals want to attract skillful students. They might achieve this by offering them high tracks within their schools, or they might also attract high-quality teacher by offering them homogenous classes. I test and discuss these channels in the next section.

Table 4: Proportion of the Raven's Variance explained by the School and Class levels.

Dependent Variable:	Proportion of Variance Explained (1)	Robust St. Errors (2)	95% C.I. Lower Bound (3)	95% C.I. Upper Bound (4)
Elementary School - Urban				
School level $Var_{s,es}/Var_{es}$	.1258	.0268	.0828	.191
Class level $Var_{c,es}/Var_{es}$	.0145	.0112	.0032	.0659
Residual	.8598	.0257	.8108	.9117
Gimnazja - Urban				
School level $Var_{s,gim}/Var_{gim}$	.2768	.1011	.1353	.5663
Class level $Var_{c,gim}/Var_{gim}$	.0936	.0294	.0505	.1733
Residual	.6297	.0502	.5386	.7362
Elementary School - Rural				
School level $Var_{s,es}/Var_{es}$	.2581	.0461	.1818	.3664
Class level $Var_{c,es}/Var_{es}$	.0135	.0079	.0043	.0423
Residual	.7284	.0298	.6722	.7893
Gimnazja - Rural				
School level $Var_{s,gim}/Var_{gim}$	.0535	.0142	.0318	.0899
Class level $Var_{c,gim}/Var_{gim}$	.06	.0156	.0361	.0997
Residual	.8865	.0333	.8236	.9543

Notes: The table shows decomposition of variance of the standardized Raven's Progressive Matrix Test Score, by the school and class level. The estimation was conducted using the mixed (hierarchical) effect model. Each stage was weighted using survey weighting scheme.

The results are different for rural areas. At the entrance to elementary school, the school and class levels explain 26% and 1% respectively of the Raven's variation. At the entrance to *gimnazjum*, the school level drops to 5%, which means that *gimnazja* are more heterogeneous than elementary schools. This is likely to be explained by the differences in catchment areas sizes. At the same time, the fraction explained by the class level rises to 6%. Interpretation of this change is less straightforward. Suppose that there is just one class per elementary school and students have exactly the same classmates in elementary school and *gimnazjum*. Because of the nested catchment areas, students from several elementary schools will go to one gimnazjum, and each class in that gimnazjum will consist of students coming from the same elementary school. As a result, the importance of the class level increases, even though there was no change in the class composition. However, this also implies that the unexplained part of variance does not change. Contrary to this, I report an increase in the unexplained part of variance, which suggests that also classes are more heterogeneous at the entrance to *gimnazjum*,

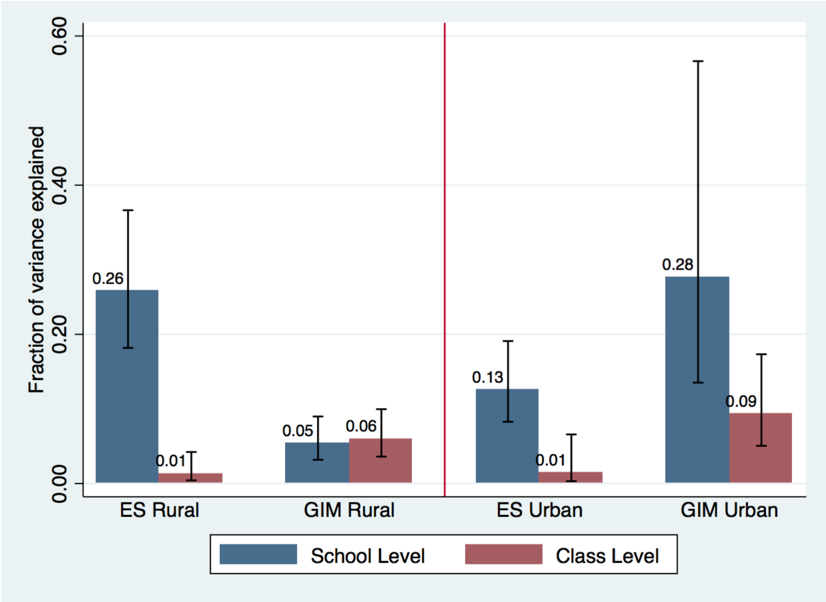


Figure 3: Proportion of Raven's Variance explained by School and Class

Note: The figure presents decomposition of variance of the standardized Raven's Progressive Matrix Score, by the school and class level using the mixed (hierarchical) effect model. Each stage was weighted using survey weighting scheme.

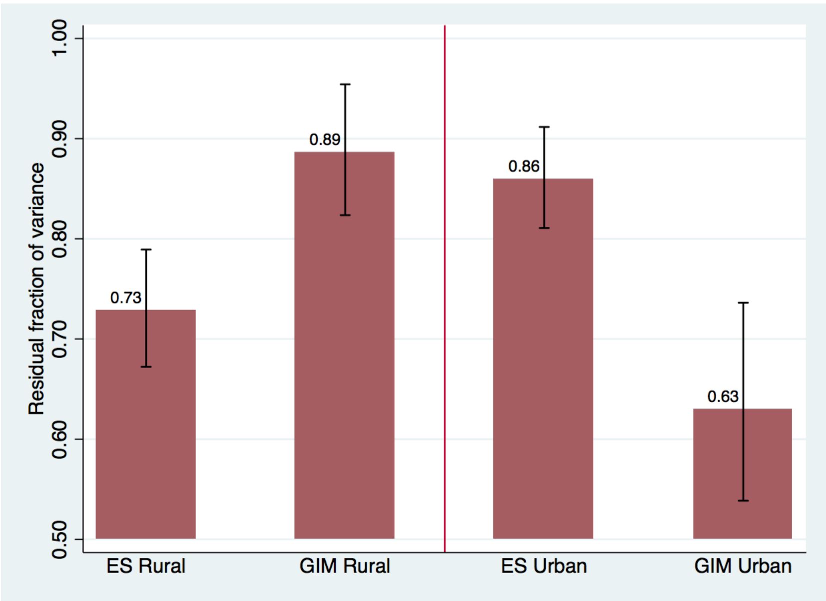


Figure 4: Residual of Raven's Variance (fraction not explained by School and Class)

Note: The figure presents unexplained (residual) proportion of variance of the standardized Raven's Progressive Matrix Score. The Estimation uses the mixed (hierarchical) effect model. The survey weighting scheme is applied.

compare to elementary school. Using Equation 4, I calculate a drop in sorting within a school to 16 pp of the fraction explained by the class level. Why would *gimazjum* principals mix incoming students across classes? The likely reason is a political pressure to fight educational inequalities, which motivates random assignment of student into classes. This possibility is discussed in more detail in the next section.

What can we learn from these numbers about the effect of school competition? For urban areas, if only Assumption 1 holds, Equation 4 provides a lower bound of the potential effect. This is because it ignores the mixing effect of catchment areas and political pressure to randomized class composition. Nevertheless, I report a 15 pp increase in the importance of the school level (ΔVar_s) and an 8pp increase in the importance of the class level (ΔVar_c). When assumptions 2, 3.a and 3.b hold, the difference between rural and urban areas provides an upper bound estimate of the school competition. For sorting between the change in the importance of the school level is $\Delta Var_s^{URBAN} - \Delta Var_s^{RURAL} = 15pp - (-21pp) = 36pp$, whereas for for sorting within the increase in the importance of the class level is $\Delta Var_c^{URBAN} - \Delta Var_c^{RURAL} = 8pp - (-16pp) = 24pp$. Table 5 summarizes these calculations.

Table 5: The Effects of Interest

		Urban (1)	Rural (2)	Difference (1)-(2)
Sort. ΔVar_c	Within	$9\% - 1\% =$ **8pp**	$6\% - 1\% + (5\% - 26\%) = -16pp$	**24pp**
Sort. ΔVar_s	Between	$28\% - 13\% =$ **15pp**	$5\% - 26\% = -21pp$	**36pp**
Interpretation		*Lower Bound*		*Upper Bound*

Notes: The table presents the logic behind the lower and upper bound estimates of the effect of school competition on sorting between schools and within a school. The numbers used in calculations come from Table 4.

Assumption 3.a says that the change in a general classroom assignment practice is the same in rural and urban areas. As argued previously, it is not restrictive and the true effect of school competition on sorting within a school should be close to the upper bound estimate (24pp). However, Assumption 3.b is unlikely to be true and the mixing effect of larger catchment should be larger in rural areas. In order to shed light on the possible magnitude of the true effect, I relax this assumption and claim that the mixing effect is proportional to the ratio of elementary schools to *gimnazja*. Table 1 shows that the ratio for rural area is 2.31 elementary schools per *gimnazjum* and for urban area the ratio is 1.49. From Table 4, in rural areas sorting between schools drops by 21pp between the two stages of education. Hence, "back of the envelope" calculations suggest that the mixing effect in urban areas is: $1.49/2.31 = 0.651$ times $21pp$, which equals $13.7pp$. Based on this, the effect of school competition on sorting between schools is $15pp + 13.7pp = 28.7pp$ of the proportion of variance explained by the school level.

5.2 Robustness

The mixed effect estimation conducted in Stata does not fully incorporate the replication weights provided by the survey authors. As a robustness check, I report the unweighted estimates in the appendix Table 9 and Figure 5. The results are sensitive to the weighting scheme. The lower (upper) bound estimate for the effect of school competition on sorting between schools is 7pp (26pp), and on sorting within a school is 14pp (26pp). Nevertheless, the direction of change is the same, except a larger importance given to sorting within a school.

Test-room shocks at the time of measurement could drive the results. Suppose that a barking dog was influencing students' attention during Raven's test. The measure of student homogeneity might by driven then not only by sorting but also by the fact that all students were exposed to the barking dog. These could explain changes between elementary schools and *gimnazja*, as in the former each student took the test separately, while in the later students took the test in groups. In particular, this would imply more homogeneous classes in *gimnazja*. There are three reasons why this scenario is unlikely. Firstly, the measurement was conducted by the team of professional psychometricians with all measures taken to provide neutral environment for all test-takers (Jasinska, Hawrot, Humenny, Majkut and Konlewski, 2013). Secondly, the nature of these shocks would have to be different between urban and rural schools. However, there are no reasons to suspect that former have more intensive test-room shocks. Thirdly, to fully exclude this possibility, I exploit the fact that in almost one-third of *gimnazja* students took Raven's test in two groups within a class. Thanks to this, I can directly check whether there is any impact of being in a separate group on Raven's score after controlling for the class fixed effects. The potential significant effect would indicate that the test-room environment matters for the outcome, however the regression shows highly insignificant coefficient, both in urban and rural areas. On the other hand, the correlation between a student's Raven's score and the average of her classmates from the same testing group is significantly higher than the correlation with the other group (from the same class). Nevertheless, the difference is larger in the rural areas which is not consistent with the test-room shock story (the results are available upon request).

The other possible explanation is difference between the cohorts in sorting at the entrance to elementary school. Specifically, for the seventh graders (from 2010) the sorting at their first grade (in 2004) could be different than for the first graders in 2010. The data do not allow me to fully explore this possibility, nevertheless I use a question on whether a child attended a local elementary school to shed light on this possibility. Table 6 presents the percentage of parents who sent their children to a local school, by the urban/rural breakdown.[11] There is a difference between the elementary school entrants in 2004 and 2010 as the seventh graders were more likely to go to their assigned schools. However, the difference is only statistically different from zero in the total sample (with p-value=4%) and the magnitudes are very small: 2.9pp for the whole sample, 1.2pp for rural schools and 4.8pp for urban. Even though this effect could possibly bias downward the results its magnitude and significance cast doubts on the importance of it.

As for sorting within a school, there are no clear reasons why principals' practice could change between 2004 and 2010. The results presented in Table 4 show that sorting within is negligible at the entrance to elementary school. Moreover, there was no institutional change which would provide additional motivation for student grouping or vice-versa. Finally, because the change in sorting within is different in rural and urban areas, the possible confounding effect would have to affect sorting in a heterogeneous way. I find this possibility rather unlikely.

To summarize, even though more data are needed to fully exclude alternative explanations, there are no convincing evidences that the main results are driven by either test-room shocks or across-cohort changes in sorting at the entrance to elementary schools.

Table 6: Percentage of students who attended a local, assigned elementary school

Stage	All	Urban	Rural	n
Elementary school	79.1%	72.4%	82.1%	7066
Gimnazjum	82%	77.2%	83.4%	4844
Difference	2.9pp**	4.8pp	1.3pp	
N	10528	3455	7073	

Notes: Percentage of answers "yes" for the question asked to parents whether their child attended a local and assigned elementary school. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level..

[11] This question is thus more retrospective for the parents of students from *gimnazjum*.

6 School Competition and Sorting Within a School

The results show that school competition leads to higher sorting of students within a school. This is in line with two theoretical works by Epple et al. (2002) and Clotfelter et al. (2005). Epple et al. (2002) argue that creation of a high track within a school might be used to attract high-skill or high-income students (the demand for peer quality), while Clotfelter et al. (2005) also suggest that it might be used to attract high-skilled teachers (the demand for teachers).

Suppose that students differ by skill levels and they maximize expected difference between benefits and costs of education. The benefits are a function of class' peer quality and teacher skills, whereas the costs depend on a school distance. Students select a non-local school only if its peers' and teachers' quality overbalance the extra costs of a longer travel distance. Next, suppose that school principals maximize enrollment and they have to accept all students coming from a local catchment area. Because the enrollment depends on expected benefits from education, principals also indirectly care about school quality. Principals decide whether to sort or mix students across classes. Sorting yields an extra cost, because principals have to adjust teaching practices and conduct selection of students. They can also hire teachers, but the offered salary is endogenously. Teachers differ by their skills and they select a school that maximizes their utility, which is an increasing function of wage (fixed across schools) and classroom environment (determined by the quality of students).

When the cost of traveling is high enough, so that students never select an alternative school, the school principals have no motivation to introduce within school tracking. This is a potential scenario for rural areas. Now, consider urban areas where school choice is feasible because the costs of traveling is low. In general, students are more likely to select a non-local school if they live in a low-quality area. In order to keep local high-achievers and attract non-local ones, school principal of a school from the low-quality area have to provide skilled teachers or a high quality class' peers. Therefore, school choice together with residential sorting, might motivate principals to use classroom sorting as a mean of competition for high-skill students (the demand for peer quality channel). Also, since teacher wages are fixed by the collective bargaining agreements, the only way for school principals to attract skilled teachers is by offering them a pleasurable teaching environment (the demand for teachers channel).

In this first part of this section I present survey data on *gimnazjum*'s principals characteristics and their sorting practices. In the remaining part of the section, I empirically evaluate the two mechanism. The results suggest that the demand for peer quality explains the positive association between school competition and class assignment.

6.1 Survey Data on School Principals

The questionnaire for *gimnazjum*'s principals might shed light on the reasons behind the increase in sorting within a school. Importantly, it asks about the class assignment procedures. The data is self-reported and has a qualitative nature thus it only provides anecdotal evidences.[12] Consistently across urban ad rural areas, the principals underline that they want to create balanced classes in terms of past GPA and the sixth-grade standardized examination scores. At that time in Poland there was a strong political pressure to equalize educational opportunities and school principals might do not want to openly speak about their sorting practices. On the other hand, the political pressure can explain why students are more mixed across classes when entering *gimnazjum* in rural areas.

The questions about the attitude toward external examinations and their usage in various school activities may be more informative about principals' behavior. Table 7 presents results for 150 *gimnazja* in the sample, Panel A shows that principals from urban schools are more likely to trust and use information coming from the external examinations, at the same time they believe that the score matters too much in an educational path of a child. These results are consistent with the observed higher sorting across classes and schools in urban areas. However, even though the magnitude is relatively large the differences are mostly insignificant. Differences in principal's characteristics might also matter. Panel B of Table 7 shows that principals have almost identical work experience[13], but the share of females is higher in urban areas.

[12]The reliability of this kind of data is discussed in Betts and Shkolnik (2000).
[13]Because of hiring criteria all principals have the same level of education.

10.14754/CEU.2016.03

Table 7: *Gimnazjum*'s Principals

Question	Urban	Rural	Difference
Panel A: Principals and the External Examination			
6th grade exam as a good signal	67.2%	55.6%	11.4pp
Usage of the 6th grade exam	84.8%	77.8%	7pp
External examination as a good signal	93.5%	83.4%	9.9pp**
External examination is random	18%	26.3%	8.3pp
External examination is too influential	62%	47%	15pp
Panel B: Principals' characteristics			
Experience in schooling (years)	24	24.3	0.3
Experience as a principal (years)	11.2	9.9	1.3
% of females	70%	60%	10pp
N	46	104	

Note: Variable "6th grade exam as a good signal" is an answer to the question "Is the 6th grade exam a good measure of skills of students who are attending your school?"; "Ext. exam as a good signal" is an answer to " Do you agree that the external examination allows to compare students' achievements?"; Ext. exam is random is an answer to: "Do you agree that the examination scores are pretty much random?"; "Ext. exam is too influential" is an answer to: "Do you agree that the examination scores matter too much in the educational path of a child?". All above variables equals one for questions:"strongly agree"/"rather agree" and 0 for "rather disagree"/"strongly disagree". Variable "Usage of the 6th grade exam" is one if principal's school analyzed examination score and used them somehow. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

6.2 The Demand for Peer Quality

The demand for peer quality channel appears when *gimnazjum* principals create a high track within a school in order to attract non-local high-skill/income students or in order to keep local ones. The available data do not allow me to investigate the later possibility, instead I empirically check whether a *gimnazjum*-level measure of student sorting based on Raven's score is correlated with sorting based on non-locality of students. If school principals create classes with high-achievers to attract non-local students, I should observe a positive association between the two measures of sorting. Since I focus on the sorting at the entrance to *gimnazjum*, I exclude observations from elementary schools. In particular, I follow Collins and Gan (2013) and define sorting of students across classes as:

$$W_s^R = \frac{1}{2}\sum_c \frac{\widehat{\sigma}_{cs}^R}{\widehat{\sigma}_s^R} \tag{7}$$

where $\widehat{\sigma}_{cs}^R$ is the observed standard deviation of Raven's score for class c from *gimnazjum* s and $\widehat{\sigma}_s^R$ is the observed standard deviation of Raven's score for *gimnazjum* s. The ratio is defined for each class within a school, but I calculate the school-level average (in the data I have two classes per school). With perfect sorting across classes, the variance for each class is zero but at the school-level is positive, hence the measure W_s^R is null. On the other hand, with perfect mixing the variance is the same for the class and school levels and W_s^R equals to one.[14] I also define a similar measure for sorting based on non-locality of students:

$$W_s^N = \frac{1}{2}\sum_c \frac{\widehat{\sigma}_{cs}^N}{\widehat{\sigma}_s^N} \tag{8}$$

where $\widehat{\sigma}_{cs}^N$ is the class-level observed standard deviation of dummy indicating whether a student is non-local and analogously $\widehat{\sigma}_s^N$ is for the school-level. My regression of interest is:

$$W_s^R = \alpha + \beta W_s^N + \epsilon_s \tag{9}$$

If the demand for peer quality channel is present I should observe a positive correlation between the two measures for sorting in urban areas and a null correlation in rural areas. Table 8 Columns (1), (3) and (5) show that switching from perfect sorting to mixing in urban areas on average increases the Raven's sorting measure by .254, which means that classes become more heterogeneous. In rural areas, the coefficient is negative and insignificant.

The measure of sorting based on non-locality of students might be misleading when there is a few non-local students. To see this suppose that there is one non-local student and he/she was randomly assigned to a class. For the assigned class the measure will be one and for the second class it will be zero. Consequently, the school-average measure of sorting will be half, even though non-local students was assigned in a random way. As an alternative measure of sorting based on non-locality I calculate the absolute difference between classes in the share of non-local students. Since there are only two classes per school, the measure is defined as $|NonLocal_{1s} - NonLocal_{2s}|$, where $NonLocal_{1s}$ is a share of non-local students in the first class from school s. The regression of interest is:

$$W_s^R = \alpha + \beta |NonLocal_{1s} - NonLocal_{2s}| + \epsilon_s \tag{10}$$

Note that larger $|NonLocal_{1s} - NonLocal_{2s}|$ implies higher sorting across classes based on non-locality of students. Consequently, the demand for peer quality channel means

[14]The sorting measure can be larger than unity. This might happen when one class consists of students from the middle of the distribution and the second class consists of students from the bottom and top of the distribution.

a negative correlation in urban areas and null in rural areas. Table 8 Columns (2), (4) and (6) show that switching from an equal share of non-local students in each class, to a situation where all non-local students are placed in one class and there are no local students with them, on average decreases Raven's Score sorting measure by -.209 (students are more sorted based on Raven's score). Again, no significant correlation is reported for rural areas.

Table 8: Channels: The Demand for Peer Quality and The Demand for Teachers

	All		Urban		Rural	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dependent Variable: Sorting based on Raven's Score (W_s^R)						
Sorting based on Non-Locality (W_s^N)	.010		.254		-.055	
	(.054)		(.070)***		(.055)	
$\lvert NonLocal_{1s} - NonLocal_{2s} \rvert$		-.096		-.209		.021
		(.046)**		(.052)***		(.047)
Constant	.948	.970	.699	.987	1.013	.959
	(.050)***	(.008)***	(.068)***	(.014)***	(.050)***	(.009)***
N	136	141	41	44	95	97
R^2	.0005	.043	.214	.268	.019	.002
Panel B Dependent Variable: Mean of Raven's score at the class level						
Teacher's Rank	-.080	-.096	-.025	-.223	-.100	-.081
	(.080)	(.101)	(.173)	(.222)	(.087)	(.111)
Teacher's Experience		.002		.026		-.003
		(.008)		(.019)		(.009)
N	267	266	90	90	177	176
R^2	.681	.682	.729	.737	.632	.635

Notes: Panel A presents results of the OLS regression of the standardized Raven's score sorting measure W_s^R on the non-locality sorting measure W_s^R and the distance in the share of non-local students between classes. The measures are calculated for the 7th graders only. The unit of observation is school (*gimnazjum*). Panel B presents the OLS regressions of the class average of the standardized Raven's score on class-averages of teacher experience in years and teacher's professional rank. The rank ranges from 0 to 5, where 0 is rank-less teacher and 5 is "the professor of education". The measures are calculated for the 7th graders only. The unit of observation is class (from *gimnazjum*). Robust standard errors are presented in the parentheses. All the estimations are weighted using the survey weights. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

This empirical evidence strongly suggest that there is a connection between school competition, sorting of students based on Raven's score and sorting based on non-locality of students. This is consistent with the demand for peer quality channel, which claims that school principals might want to attract non-local students by offering them high-tracks. Nevertheless, more research is needed to investigate the possibility that school principals use high-tracks to keep local students in their local schools.

6.3 The Demand for Teachers

School principals might want to compete for students by offering them high quality teachers. However, in many countries (including Poland) teacher wages are set through the teacher collective bargaining agreements and cannot be modified by school principals. Assuming that teachers prefer homogeneous and high-achieving classes, principals could attract skilled teacher by sorting students across classes. To test this possibility, I correlate teacher characteristics with the class average of Raven's score and control for the school fixed effects. If the demand for teachers mechanism is present, I should observe a positive association between the measures of teacher experience and the class-average of Raven's score. I focus only on teachers and classes from *gimnazja*. The regression of interest is specified as follows:

$$\overline{Y}_{cs} = \alpha + \beta T_{cs} + \mu_s + \epsilon_{cs} \tag{11}$$

where $\overline{Y}_{cs}$ is the average Raven's Score for class c from *gimazjum* s, T_{cs} is the class-average of teacher characteristics and μ_s are the school fixed effects. I use two measures of teaching characteristics: professional ranks and teaching experience in years. There are five ranks ranging from rankless teacher (=0) to "the professor of education" (=5).

Table 8 Columns (1), (3) and (5) present the correlations between the class-averages of teacher's rank and Raven's score. The correlations are negative but insignificant across samples, with the magnitudes larger in urban areas. Columns (2), (4) and (6) show the same regression, but with teaching experience as an additional independent variable. The magnitude of the coefficient on teacher's rank doubles for the urban sample, but still remains insignificant. Similarly, the coefficient on teaching experience is practically null.

The results suggest that *gimnazjum* principals do not offer high-tracks in order to attract skilled teachers. Interestingly, the results, even though imprecise, suggest that school principals might assign higher-rank teachers to worse classes in urban areas, but not in rural. They might want to compensate lower-peer quality with better teachers or handling lower tracks require teachers with higher skills.[15] Regardless of the reasons, this might have a positive effect on educational equality of opportunity. More data is needed to fully investigate this possibility.

[15] An alternative, but somehow unlikely, explanation is that high-skill teachers prefer to teach low-skill students. But in this scenario, high-skilled students will be not attracted by teacher quality and thus school principals have no motivation to hire skilled teachers.

7 Conclusions

In this study I estimate the effect of school competition on sorting within a school (across classes). The identification strategy is based on a two-stage design of the Polish Comprehensive Education, which allows me to isolate an exogenous change in student mobility. In addition, I use a novel measure of student socio-economic characteristics - Raven's Progressive Matrix test score. The results show that school competition leads to a higher sorting of students within a school and between schools. I investigate two theoretical explanation of the effect on sorting within a school: the demand for peer quality (Epple et al., 2002) and the demand for teachers (Clotfelter et al., 2005). The data point to the importance of the former mechanism, i.e. the demand for high quality peers that motivates school principals to create high tracks within a school.

This paper might be useful for policymakers who wish to use school competition as a mean to improve quality of schools, but also wish to avoid its negative distributional consequences. The results underlines the importance of school principals' incentive structure. Classroom assignment, by creating classes with high *level* of peer quality, might be used by principals to attract high-achievers or high-income students. This could be weaken by the incorporation of value added estimates of school performance into principals' objectives, as it motivates them to compete also for low-background or low-performing students (MacLeod and Urquiola, 2009). Even though the value-added based accountability has been heavily discussed, not much attention has been paid to the potential distributional effects (Rothstein, 2009; Angrist, Pathak and Walters, 2011; Chetty, Friedman and Rockoff, 2014). Alternative policy could be to link school vouchers with the socioeconomic background, for instance to offer them only to students with low income [16] On the other hand, abolishing the teacher collective bargaining agreements allows school principals to compete based on wages rather than composition of students. Nevertheless, in this study I do not find strong evidence for the demand for teachers mechanism.

References

Angrist, Joshua D, Parag A Pathak, and Christopher R Walters, "Explaining charter school effectiveness," Technical Report, National Bureau of Economic Research 2011.

[16] This policy is in effect in Chile and the Netherlands, see Böhlmark et al. (2015).

Betts, Julian R and Jamie L Shkolnik, "Key difficulties in identifying the effects of ability grouping on student achievement," *Economics of Education Review*, 2000, *19* (1), 21–26.

Böhlmark, Anders, Helena Holmlund, Mikael Lindahl et al., "School choice and segregation: evidence from Sweden," Technical Report 2015.

Brouwers, Symen A, Fons JR Van de Vijver, and Dianne A Van Hemert, "Variation in Raven's Progressive Matrices scores across time and place," *Learning and Individual Differences*, 2009, *19* (3), 330–338.

Chetty, Raj, John N Friedman, and Jonah E Rockoff, "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates," *The American Economic Review*, 2014, *104* (9), 2593–2632.

Clotfelter, Charles T, Helen F Ladd, and Jacob Vigdor, "Who teaches whom? Race and the distribution of novice teachers," *Economics of Education review*, 2005, *24* (4), 377–392.

Collins, Courtney A and Li Gan, "Does sorting students improve scores? An analysis of class composition," Technical Report, National Bureau of Economic Research 2013.

Dolata, Roman, *Szkoła, segregacje, nierówności*, Wydawnictwa Uniwersytetu Warszawskiego, 2008.

Epple, Dennis and Richard E Romano, "Competition between private and public schools, vouchers, and peer-group effects," *American Economic Review*, 1998, pp. 33–62.

_ , Elizabeth Newlon, and Richard Romano, "Ability tracking, school competition, and the distribution of educational benefits," *Journal of Public Economics*, 2002, *83* (1), 1–48.

Figlio, David N and Marianne E Page, "School choice and the distributional effects of ability tracking: does separation increase inequality?," *Journal of Urban Economics*, 2002, *51* (3), 497–514.

_ and Maurice E Lucas, "What's in a grade? School report cards and the housing market," *American Economic Review*, 2004, pp. 591–604.

Friedman, Milton, *The role of government in education*, Rutgers University Press, 1955.

Gibbons, Stephen, Stephen Machin, and Olmo Silva, "Choice, competition, and pupil achievement," *Journal of the European Economic Association*, 2008, *6* (4), 912–947.

Herbst, Mikołaj, Jan Herczyński, and Anthony Levitas, *Finansowanie oświaty w Polsce: diagnoza, dylematy, możliwości*, Wydawnictwo Naukowe Scholar, 2009.

Hoxby, Caroline M, "Does competition among public schools benefit students and taxpayers?," *American Economic Review*, 2000, pp. 1209–1238.

Hsieh, Chang-Tai and Miguel Urquiola, "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of Public Economics*, 2006, *90* (8), 1477–1503.

Instytut Badań Edukacyjnych, "Raport o stanie edukacji 2010," 2011.

Jasinska, Aleksandra, Joanna Hawrot, Grzegorz Humenny, Przemyslaw Majkut, and Maciej Konlewski, "Kontekstowy model efektywnosci nauczania po I etapie edukacyjnym - wyniki badania podluznego w szkolach podstawowych.," Technical Report, Instytut Badan Edukacyjnych 2013.

Jensen, Arthur Robert, *The G Factor: The Science of Mental Ability*, Praeger, 1998.

Kalogrides, Demetra, Susanna Loeb, and Tara Béteille, "Systematic sorting teacher characteristics and class assignments," *Sociology of Education*, 2013, *86* (2), 103–123.

Kane, Thomas J, Stephanie K Riegg, and Douglas O Staiger, "School quality, neighborhoods, and housing prices," *American Law and Economics Review*, 2006, *8* (2), 183–212.

Kern, Nora, Renita Thukral, and Todd Ziebarth, "A Mission to Serve: How Public Charter Schools Are Designed to Meet the Diverse Demands of Our Communities. Issue Brief.," *National Alliance for Public Charter Schools*, 2012.

Kremer, Michael, Esther Duflo, and Pascaline Dupas, "Peer Effects, Teacher Incentives, and the Impact of Tracking," *American Economic Review*, 2011, *101* (5 (August 2011)).

Ladd, Helen and Edward Fiske, "When schools compete: A cautionary tale," *Washington, DC: The Brookings Institution*, 2000.

MacLeod, W Bentley and Miguel Urquiola, "Anti-lemons: school reputation and educational quality," Technical Report, National Bureau of Economic Research 2009.

Meghir, Costas and Mårten Palme, "Educational reform, ability, and family background," *The American Economic Review*, 2005, *95* (1), 414–424.

Nechyba, Thomas J, "Income and peer quality sorting in public and private schools," *Handbook of the Economics of Education*, 2006, *2*, 1327–1368.

Raven, John, "The Raven's progressive matrices: change and stability over culture and time," *Cognitive Psychology*, 2000, *41* (1), 1–48.

Rothstein, Jesse, "Student sorting and bias in value-added estimation: Selection on observables and unobservables," *Education*, 2009, *4* (4), 537–571.

Sochocki, Marcin, "Raport: Szkoly i przedszkola na wsi i malych miasteczkach.," *Federacja Inicjatyw Oświatowych*, 2011.

Szmigel, Maria Krystyna, "Gimnazjum niejedno ma imię? - na podstawie wywiadoów z dyrektorami. Doniesienie z badań jakościowych," Technical Report, XIX Konferencja Diagnostyki Edukacyjnej, Gniezno 2013 2013.

Tiebout, Charles M, "A pure theory of local expenditures," *The Journal of Political Economy*, 1956, pp. 416–424.

Appendix

Table 9: Proportion of Raven's Variance explained by School and Class - unweighted estimates

Dependent Variable:	Proportion of Variance Explained (1)	Robust St. Errors (2)	95% C.I. Lower Bound (3)	95% C.I. Upper Bound (4)
Elementary School - Urban				
School level $\sigma_{s,es}/\sigma_{TOT,es}$	0974	.0262	.0575	.165
Class level $\sigma_{c,es}/\sigma_{TOT,es}$	.0233	.0139	.0073	.075
Residual	.8793	.0279	.8262	.9358
Gimnazja - Urban				
School level $\sigma_{s,gim}/\sigma_{TOT,gim}$	.1651	.0619	.0792	.3441
Class level $\sigma_{c,gim}/\sigma_{TOT,gim}$	.161	.0441	.0941	.2754
Residual	.6739	.0252	.6263	.7251
Elementary School - Rural				
School level $\sigma_{s,es}/\sigma_{TOT,es}$	.2106	.034	.1534	.2891
Class level $\sigma_{c,es}/\sigma_{TOT,es}$	.0205	.0103	.0076	.0551
Residual	.769	.0189	.7327	.807
Gimnazja - Rural				
School level $\sigma_{s,gim}/\sigma_{TOT,gim}$	.0189	.0159	.0036	.0987
Class level $\sigma_{c,gim}/\sigma_{TOT,gim}$	.086	.02	.0545	.1358
Residual	.8951	.0225	.8521	.9402

Notes: The table shows decomposition of variance of the standardized Raven's Progressive Matrix Score, by the school and class level. The estimation was conducted using the mixed (hierarchical) effect model.

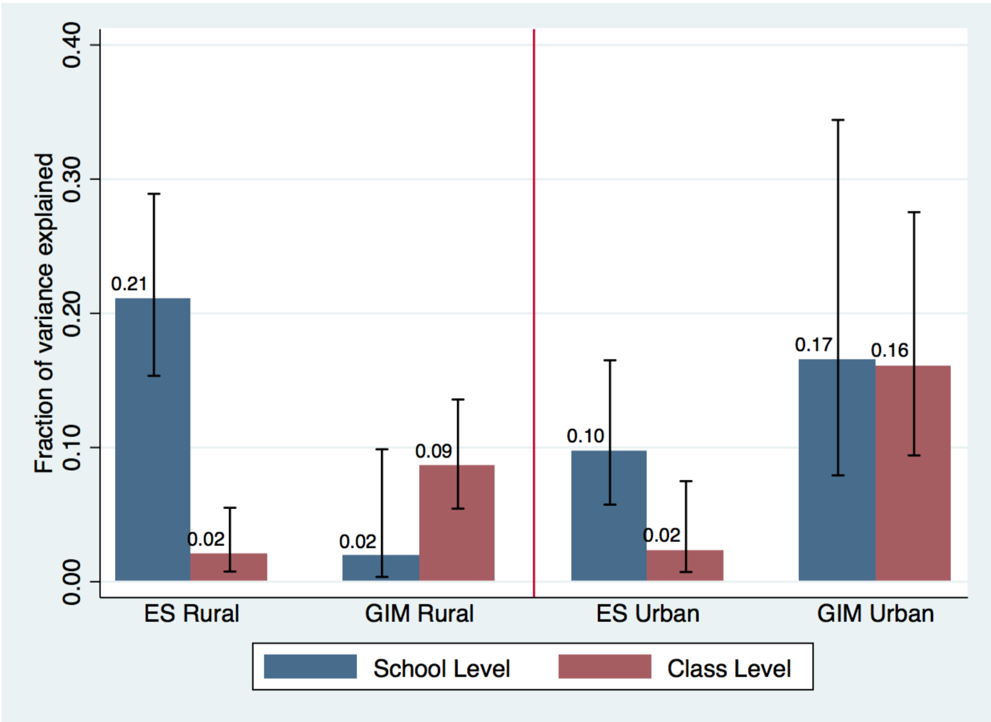


Figure 5: Proportion of Raven's Variance explained by School and Class - Unweighted Estimation

Note: The figure presents decomposition of variance of the standardized Raven's Progressive Matrix Score, by the school and class level using the mixed (hierarchical) effect model.