Essays on Development Economics and Economic History

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2017

vorgelegt von

Andreas Backhaus

Referent: Prof. Dr. Uwe Sunde
Korreferent: Prof. Dr. Ludger Wößmann

Promotionsabschlussberatung: 31. Januar 2018

Datum der mündlichen Prüfung: 18.01.2018

Namen der Berichterstatter: Uwe Sunde, Ludger Wößmann, Matteo Cervellati

Acknowledgements

First, I would like to thank Stephan Klasen and Gilles Saint-Paul, who recommended me for the PhD position in Munich. I am also much obliged to Dalia Marin, who made the initial decision to accept me into the MGSE.

Next, I would like to thank my primary supervisor Uwe Sunde, who accepted me as a member into the Chair for Population Economics. I particularly appreciate that he always encouraged every new line of my research without imposing how I pursue it. Being granted the freedom to try out and to learn was elementary to my doctoral studies that led up to this dissertation. I would further like to thank Ludger Wößmann, who agreed to act as my secondary supervisor, and Matteo Cervellati for serving as the oral examiner on my dissertation committee. Eva Tehua and Susan Fay were vital for mastering all administrative procedures in the background.

I am grateful for the opportunity to have worked with Inma Martinez-Zarzoso and Chris Muris on our joint paper. I would further like to express my gratitude towards Jan-Otmar Hesse, who kept my interest in the history of economic thought alive.

My fellow PhD students, in particular at the Chair of International Economics and the Chair for Population Economics, contributed significantly to my endurance and my motivation over the course of the years. In particular, I would like to thank my officemates Gerrit and Rainer, who made this episode much more fun than it would have been without them. Dina always encouraged both my interest in Central and Eastern Europe and my curiosity about hiking, both of which turned out to be beneficial for my PhD advancement and lifestyle. Rudi Stracke and Lukas Buchheim were always open to share their experiences with me over a seemingly inelastic supply of coffee.

I would like to thank the people who continuously supported me simply because they cared, among them Jana, Kathrin, Marina, Martin, Nastya, and Richard.

Last but not least, thanks to all the people back home who stayed with me and who did not grow tired of backing me up! A special word of thanks goes to my mother for her unfailing faith and encouragement.

Andreas Backhaus, May 2018

Contents

List of Figures

List of Tables

Preface

Over the recent years, the research domains of development economics and economic history have both witnessed significant augmentations of their empirical toolsets to estimate causal relationships. Differences-in-differences, instrumental variable methods and regression discontinuity designs (RDD) have now become standard techniques as a consequence of what some decided to call "the credibility revolution in empirical economics" (Angrist & Pischke, 2010). At around the same time, both research domains have also benefited from unprecedented expansions of their available data. While economic historians have tapped into the abundant reservoirs of historical maps and statistics, e.g. on education (Becker & Woessmann, 2009; Becker et al., 2014) or income (Piketty, 2003), development economists have been referred to the potential of remote sensing data, in particular nighttime lights, by contributions such as Henderson et al. (2012). Unsurprisingly, fruitful overlappings between the two domains have since then followed (e.g. Michalopoulos & Papaioannou, 2013a; Michalopoulos & Papaioannou, 2013b).
The recent advancements in terms of both empirical refinement and data provision manifest themselves in the three contributions to development economics and economic history that form the main body of this thesis. Each contribution corresponds to one chapter; all chapters are self-contained and can be read independently from each other.
It is the importance of history for the formation of human capital that is highlighted by the first chapter of the thesis. A large-scale natural experiment had instilled the country of Poland with three different legacies of education. Known as the Partitions of Poland (1795-1918), the Polish nation was divided among the empires of Austria, Prussia and Russia, resulting in sharply diverging economic and social conditions across the partitions by the time they were merged into a common Polish state again. While Grosfeld & Zhuravskaya (2015) show that the past economic and educational disparities cannot be detected along the former partition borders in present-day Poland anymore, I intend to decompress the Polish history that led up to this important result in the spirit of Austin (2008). The chapter is further greatly inspired by the notion of Wittenberg (2015) that a legacy from the past is as much a product of history as a non-legacy, as *any* conceivable outcome must necessarily result from *some* prior causal factor.

Extensive data collection allows me to construct a georeferenced sample on historical school enrollment, school supply and basic human capital covering the partition territories from the times of the empires to the height of communism in Poland. Using the exogeneity of the partition borders for a spatial RDD along the Austrian-Russian and Prussian-Russian borders, I find that the causal effect of the partitions on primary school enrollment is initially large and unfavorable for the partition under Russian rule shortly before the outbreak of World War I (WWI). However, I show that this effect disappears within the following two decades of reconstituted Polish independence. The disappearance is permanent, as there is no evidence for a rebound during the communist era after World War II (WWII). Similarly, the backlog of the Russian partition with regard to literacy declines over time, albeit at a slower rate than schooling, reaching negligible importance in the 1960s. While my investigation suggests that changes in the population composition of Poland were of only minor importance, I provide complementary evidence that the fading imperial legacies on education were accompanied by a substantial expansion of the supply of educational facilities in terms of schools and their endowment.

The works that are most closely related to this chapter are Grosfeld et al. (2015), Bukowski (2016), and Wysokinska (2017), who also make use of the partition borders in spatial RDDs. Bukowski (2016) finds a legacy of the Austrian-Hungarian Empire on education in the form of higher student test performance today on the Austrian side of the former border. Complementary evidence suggests that it was transmitted through positive social norms towards education resulting from the Austrian rule. This hypothesis does not conflict with my results of a rapidly fading legacy, as I focus exclusively on the extensive margin of education provision, while student performance relates to the intensive margin of educational quality. The particular setting of the Partitions of Poland is also the topic of Wolf (2005) and Trenkler & Wolf (2005), who study the economic integration of the Polish territories between WWI and WWII. The chapter is further closely related to Dupraz (2017) who, while in the context of (post-)colonial Africa, also studies the disappearance of a historical legacy in education along an arbitrarily drawn border.

From a methodological perspective, this chapter validates and strengthens the approach of using the past borders of the Partitions of Poland for spatial RDDs in order to estimate causal effects of the three empires. Specifically, I test whether considerable geographical discontinuities at the former borders, which have already been detected by Bukowski (2016), translate into discontinuities in agricultural suitability. The latter might then exert differential effects on the occupational choice and the returns to education on either side of a partition border. This would hint at a potential invalidity of the spatial RDD as it would be the discontinuity in agricultural suitability that is driving the discontinuities in education instead of the empires. I make use of two well-established measures of agricultural suitability (Galor & Özak, 2016) and historical croplands (Ramankutty et al., 1999). Out of the two, the suitability measure is indeed significantly discontinuous at the partition borders. However, I show that the discontinuity completely disappears at the Austrian-Russian border and shrinks to a minuscule size at the Prussian-Russian border once the discontinuous geographical variables are added as controls to the regressions, as suggested by Bukowski (2016). Consequently, all regression results reported in this chapter follow this practice.

To the best of my knowledge, this chapter represents also the first contribution that utilizes data that predate the partitions in order to perform a rudimentary check of the absence of pretreatment discontinuities at the borders. So far, the literature has relied on the assessment of historians that the demarcations of the partitions by the three imperial powers were not based on any preexisting demographic, economic or religious conditions, but rather followed military considerations or geographic features such as rivers. I complement this assessment with quantitative evidence from a population census conducted in the Duchy of Warsaw in 1810, i.e. before the finalization of the partition borders in 1815. After having georeferenced the data originally published by Grossman (1925), my estimates show no significant discontinuities in (log) population at the future partition borders. While population is arguably a crude correlate of prosperity, its utilization is not uncommon in the literature, which is why I cautiously interpret the absence of population discontinuities in the early 19th century as reassuring for the validity of the spatial RDD at the partition borders.

More generally, this chapter contributes to the literature on the economic history of education in Europe (Becker & Woessmann, 2009; Becker et al., 2011; Cantoni & Yuchtman, 2013; Squicciarini & Voigtländer, 2015; Dittmar & Meisenzahl, 2016). Further, it appends to the literature on the economic history of Central and Eastern Europe, with Grosfeld et al. (2013) and Markevich et al. (2017) in the context of the Russian Empire, while Becker et al. (2016) focus on legacies of the Austrian-Hungarian Empire.

The second chapter of the thesis is concerned with the role of ethnic favoritism in the highly competitive and ethnically diverse democracy of Ghana. It is motivated by the attention that the field of development economics currently attributes to the phenomenon of ethnic favoritism. Needless to say, investigations into ethnic favoritism have been largely directed at countries and continents where ethnic identities are considered to be strong and relevant. For example, Burgess et al. (2015) show that ethnic favoritism, as measured in terms of road investments in Kenyan districts that are co-ethnic with the respective political ruler, has been widespread under various autocratic rulers of different ethnic affiliation, while it appears to have been curbed after Kenya's transition to democracy in the early 1990s. Kramon & Posner (2016), in turn, assess that with regard to education, ethnic favoritism in Kenya has all but disappeared under democracy. In the broader context of sub-Saharan Africa, Franck & Rainer (2012) estimate large and widespread ethnic gains in terms of education and health originating from time periods when an ethnic group has been co-ethnic with its respective country's leader. On average, these gains remain unaffected by whether the form of government is democratic or autocratic. De Luca et al. (2016) document widespread favoritism in terms of nighttime lights on the level of ethnic homelands. Using a global dataset, they suggest that ethnic favoritism is common across, but not limited to sub-Saharan Africa, with rather negligible dampening effects of the quality of political institutions.

Under which circumstances could ethnic favoritism actually occur in a democratic system of governance? As pointed out by Amodio & Chiovelli (2016), democracy can broaden the scope for strategic interactions between politicians and ethnic leaders. The mixed evidence on the prevalence of ethnic favoritism under democracy suggests that the outcomes of these interactions might be heterogeneous with regard to the extent of ethnic favoritism that they are able to provide (or prevent). From the perspective of economic theories of democracy, the empirical prevalence of ethnic favoritism under democracy can be related to the 'core' voter concept of Cox & McCubbins (1986): Political parties differ with regard to their ability to redistribute towards different groups among the electorate, while the groups in turn differ strongly in terms of their ideological party preference. In equilibrium, this results in each party focussing its redistributive efforts on the specific group(s) to which it can redistribute the easiest. Hence, groups are generally not courted by more than one party, making them solid, delimited political blocks which enjoy high patronage as soon as 'their' party climbs to power.

At first sight, the case of the West African country of Ghana bears a large similarity to other nations whose democratization has been the subject of previous studies: Ghana returned to constitutional democracy in 1992 in the course of the so-called third wave of democratization (Huntington, 1991) after decades of alternating democratic turmoil and military rule. The latter culminated in the eleven-year reign of Jerry John Rawlings as the Chairman of the Provisional National Defence Council. Similar to the Kenyan experience, Rawlings then formed a political party, the New Democratic Congress (NDC), and went on to become the first democratically elected president of the Ghanaian fourth republic. All quadrennial elections since 1992 have been considered free and fair; they have also resulted in the first peaceful transition of power between the NDC and its main challenger, the New Patriotic Party (NPP), in the year 2000 when the constitution barred Rawlings from running for a third presidential term. From the viewpoint of ethnicity, Ghana's ethnic landscape is as much a product of the arbitrariness of colonial borders as many other African nations. The Asante, who have once been the rulers of the powerful

Ashanti empire, exhibit a strong and long-lived ethnic affiliation to the NPP. The NDC, in turn, has strong ties to the Ewe ethnic group, given that Jerry John Rawlings belongs to this group as well.

However, it is questionable whether Ghana's ethnic setup features the necessary preconditions for large-scale redistribution towards the respective president's co-ethnics because none of the two politically invested groups is actually large enough to secure a majority of the national vote by its own. This suggests that the two political parties rather have to compete for the votes of the more unaffiliated Ghanaian electorate outside of the parties' ethnic boundaries. This political constellation bears more similarity to the probabilistic or 'swing' voting models pioneered by Lindbeck & Weibull (1987) and Dixit & Londregan (1996). The essential prediction of these models is that those groups that contain the highest share of non-partisan, 'moderate' voters will be promised the highest share of redistributive transfers by the political parties because the moderates are the most likely to 'swing' their votes from one party to the other in return for economic remuneration. Correspondingly, survey evidence in Lindberg & Morrison (2005) and Lindberg (2012) indicates a relatively high and growing share of swing voters among the Ghanaian electorate. Evidence by Banful (2011) on the political economy of local budget allocations by the Ghanaian central government further suggests that districts which exhibit tighter vote margins in a presidential election receive higher allocations afterwards. Banful (2011) does not explore any ethnic dimensions of these voting patterns though.

This chapter attempts to broaden the scope for the understanding of ethnic favoritism. It does so by exploiting electoral results and changes of government in Ghana between 1992 and 2008. I first show that the two Ghanaian ethnicities that are co-ethnic with the varying presidents become economically *worse off* in relative terms soon after the country's return to democracy. I then test the prediction of the probabilistic voting theory that close voting should be associated with economic transfers if there are moderate groups of voters to be swayed. I find that there is indeed a positive association between close voting and economic prosperity, thereby confirming the finding of Banful (2011). However, I further show that this association runs entirely through the homeland of the large, politically unaffiliated ethnic group of the Akan, while it is not detectable with regard to other ethnicities. Taken together, these results suggest that while the eagerness of political parties to form multi-ethnic electoral coalitions has the effect of constraining ethnic favoritism towards the co-ethnics of the respective political leader, the same eagerness can give rise to ethnic favoritism directed towards groups that signal their readiness to be courted by the political contestants.

The chapter hence contributes to the aforementioned literature on ethnic favoritism in democracies. It further makes a methodological contribution to the literature: Because there are no official statistics on economic prosperity at the district level, which is the level of observation in the following, I make use of nighttime lights as a proxy variable. While nighttime lights have already been extensively utilized in the context of ethnic favoritism, I reaffirm the usefulness of nighttime lights for detecting patterns of favoritism at the sub-national level.

Nighttime lights as one type of remote sensing data are also central to the third and last contribution of this thesis. Their applications to economics in general and to development economics in particular have been steadily growing over the recent years (Donaldson & Storeygard, 2016; Huang et al., 2014), with the second chapter of this thesis apparently being no deviation from the trend.

Therefore, the third chapter takes a systematic approach to evaluate the applicability and consistency of nighttime lights in the development nexus. Although it is considered good practice to test the applicability of nighttime lights to the respective research context, mostly by correlating them against the variables to be proxied, the settings are so different and numerous that they are often not directly comparable with each other. This makes it difficult to derive general conclusions about the applicability and consistency of nighttime lights from them. Therefore, the approach of this chapter attempts to hold as many elements of the empirical framework fixed as possible while switching single parameters on and off one after another. In order to do so, I construct a spatially harmonized dataset from IPUMS (2017) census extracts. It allows me to examine the behavior of nighttime lights along two important dimensions: Firstly, the level of spatial aggregation can be shifted between the regional and the more disaggregated district level. Secondly, the correlations can be estimated either by pooling all observations or by exploiting the panel structure of the data. While the four resulting combinations are by far not exhaustive of the potential of nighttime lights, they still provide some clean-cut evidence on the questions whether nighttime lights can be discretionarily utilized in different spatial frameworks without loss of consistency and whether nighttime lights correlate with variables in levels as well as with their changes over time. Thereby, the chapter contributes to the literature that intends to establish nighttime lights as a valid proxy variable in many different applications (Henderson et al., 2012; Chen & Nordhaus, 2011; Chen & Nordhaus, 2015; Hodler & Raschky, 2014; Michalopoulos & Papaioannou, 2013).

The findings presented in this chapter show that firstly, nighttime lights correlate strongly with both the electrification rate of households and the average years of schooling as two consistently measured outcomes in my dataset. The correlations confirm that, within my framework, the applicability of nighttime lights as a proxy clearly goes beyond variables which feature a somewhat direct relationship to the emittance of lights such as electrification or industrial production, as the association of nighttime lights and schooling is

equally strong in terms of statistical significance and also linear in shape.
Secondly, shifting the level of spatial aggregation from the regional to the district level does not impair the ability of nighttime lights to correlate with the two outcome variables. While the district level is far from exhaustive with regard to the possibilities of disaggregating nighttime lights (or other remote sensing data), the district is already a relatively small administrative entity. This suggests that nighttime lights bear the potential to be usefully applied also within individual countries if the number of disaggregated administrative units is large enough in order to provide sufficient spatial variation.
Thirdly, nighttime lights remain strongly correlated with electrification and schooling even if a large amount of variation is discarded by estimating fixed-effects models and identifying the parameter of nighttime lights only from within-region or within-district variation. This suggests that nighttime lights are not just a stationary spatial feature, but that they co-evolve with the underlying economic, demographic and social patterns.
Taken as a whole, the results of this chapter suggest that nighttime lights maintain their consistency as a proxy variable when varying the parameters of the application systematically along the spatial and temporal dimension respectively. The assessment that nighttime lights correlate strongly and quite unconditionally with the years of schooling as a measure of educational attainment might be interpreted as particularly encouraging in the context of developing countries where disaggregated and consecutive educational data are at least as rare as GDP-related data. The correlation between nighttime lights and schooling further calls for replacing schooling with rather quality- and skills-oriented measures of human capital which Hanushek & Woessmann (2012a) and Hanushek & Woessmann (2012b) deem as much more relevant and informative than the mere length of education. In addition, upcoming new datasets of nighttime lights hold out the prospect of even better proxy abilities, which remain to be tested by future research.

Chapter 1

Fading Legacies: Human Capital in the Aftermath of the Partitions of Poland

"Dla chcącego nic trudnego." For the willing, nothing is difficult. (Polish proverb)

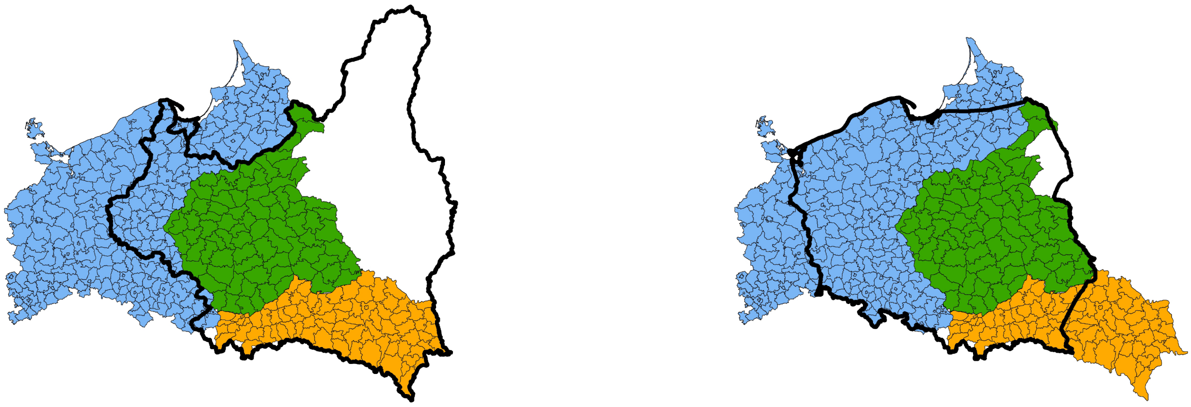
1.1 Introduction

The consensus of the empirical literature on economic history is that history matters (see Nunn, 2009, for a review). The persistence of history and thereby the importance of historical legacies have been documented for institutions (Acemoglu et al., 2001; Michalopoulos & Papaioannou, 2013), human capital (Glaeser et al., 2004; Valencia, 2015), and technology (Comin et al., 2010), to name only a few examples.

Indeed, it would actually be rather surprising if history did not matter: As Wittenberg (2015) points out, *any* conceivable outcome must necessarily result from *some* prior causal factor. From this perspective, a legacy from the past is as much a product of history as a non-legacy. Considering, for example, on a large scale the fact that basic education has virtually become universal in all western European nations over the course of the 19th century, it should be evident that history mattered tremendously by turning more than a millennium of persistent human illiteracy into a non-legacy.

In this paper, I highlight the importance of history for the formation of human capital in the aftermath of a large-scale natural experiment that had instilled the country of Poland with three different legacies of education. The Partitions of Poland (1795-1918) divided the Polish nation among the empires of Austria, Prussia and Russia, resulting in sharply diverging economic and social conditions across the partitions by the time they were merged into a common Polish state again. Figure 1.1 displays the partition territories (a) within the national boundaries of the Polish state between 1919 and 1939 and (b) within the boundaries of Poland since 1944.

Figure 1.1: Partition territories within national boundaries of Poland



(a) Partition counties in interwar Poland

(b) Partition counties in Poland since 1944

While Grosfeld & Zhuravskaya (2015) show that the past economic and educational disparities cannot be detected along the former partition borders in present-day Poland anymore, I intend to decompress the Polish history that led up to this important result in the spirit of Austin (2008).

Extensive data collection allows me to construct a georeferenced sample on historical school enrollment, school supply and basic human capital covering the partition territories from the times of the empires to the height of communism in Poland. Using the exogeneity of the partition borders for a spatial regression discontinuity design (RDD) along the Austrian-Russian and Prussian-Russian borders, I find that the causal effect of the partitions on primary school enrollment is initially large and unfavorable regarding the partition under Russian rule shortly before the outbreak of World War I (WWI). However, I show that the effect disappears within the following two decades of reconstituted Polish independence. The disappearance is permanent, as there is no evidence for a rebound during the communist era after World War II (WWII). Similarly, the backlog of the Russian partition with regard to literacy declines over time, albeit at a slower rate than schooling, reaching negligible importance in the 1960s. While my investigation suggests that changes in the population composition of Poland were of only minor importance, I provide complementary evidence that the fading imperial legacies on education were accompanied by a substantial expansion of the supply of educational facilities in terms of schools and their endowment.

The works that are most closely related to my study are Grosfeld et al. (2015), Bukowski (2016), and Wysokinska (2017), who also make use of the partition borders in spatial RDDs. Bukowski (2016) finds a legacy of the Austrian-Hungarian Empire on education in the form of higher student test performance today on the Austrian side of the former border. Complementary evidence suggests that it was transmitted through positive social norms towards education resulting from the Austrian rule. This hypothesis does not conflict with my results of a rapidly fading legacy, as I focus exclusively on the extensive margin of education, while student performance relates to the intensive margin. The particular setting of the Partitions of Poland is also the topic of Wolf (2005) and Trenkler & Wolf (2005), who study the economic integration of the Polish territories between WWI and WWII. My paper is further closely related to Dupraz (2017) who, while in the context of (post-)colonial Africa, also studies the disappearance of a historical legacy in education along an arbitrarily drawn border.

More generally, this paper contributes to the literature on the economic history of education in Europe (Becker & Woessmann, 2009; Becker et al., 2011; Cantoni & Yuchtman, 2013; Squicciarini & Voigtländer, 2015; Dittmar & Meisenzahl, 2016). Further, it appends to the literature on the economic history of Central and Eastern Europe, with Grosfeld et al. (2013) and Markevich et al. (2017) in the context of the Russian Empire, while Becker et al. (2016) focus on legacies of the Austrian-Hungarian Empire.

The remainder of the paper is structured as follows: In Section 1.2, I outline the historical context of Poland in terms of population and education over the timespan of my sample. Next, I describe the available data in Section 1.3. Section 1.4 comprises my identification strategy, its assumptions and a discussion whether the latter are likely to hold within the framework of the Partitions of Poland. Section 1.5 presents my main results on the causal effect of the partitions on primary school enrollment and literacy in Poland over the period 1911-1960, followed by a concise series of robustness checks. I discuss the mechanisms behind my findings in Section 1.6. Section 1.7 concludes.

1.2 Historical context

The Partitions of Poland represented one of many drastic watersheds in Polish history. Between 1795 and 1918, Poland did not exist as an independent state, as the three neighboring powers of the Polish-Lithuanian Commonwealth, namely Prussia, Austria and Russia, had divided the territory of the large, but weakening Commonwealth among each other. Only at the end of WWI, when the Austrian-Hungarian Empire had fallen apart, the Russian Empire had descended into civil war, and the German Empire had been forced to declare a truce, Polish independence was restored in form of the Second Republic of Poland[1], reflecting the sustained Polish desire for national unity.

However, Poland could hardly be called a unified nation in 1918. Its borders in the west

[1] While it was in fact an elective monarchy, the Polish-Lithuanian Commonwealth has been called *Rzeczpospolita* (Republic) in Polish. Therefore, the reign between WWI and WWII has been attributed to the *Second* Republic.

were still to be finalized by the Treaty of Versailles, while its eastern borders were soon to be redrawn by the outcome of the Polish-Soviet war in 1920. It was also still deeply divided within, as the three empires had not failed to leave their marks on the Polish lands. By the time the borders between them were lifted, the partitions differed substantially along several dimensions, both materially and less visibly in terms of culture and institutions. At the brink of their independence, the Poles therefore saw themselves confronted with the task of consolidating the three former partitions into one national state. This challenge was particularly pronounced in the realm of education, as the empires had cared to very different extents about setting up an actual educational system: Primary schooling had already been obligatory in Prussia and Austria-Hungary for decades, but it had never even been established in the Russian Empire, creating a large disparity in educational attainment between the populations of the partitions. Consequently, addressing this disparity became the prime goal of educational policy in the following decades.
The remarks in this section serve two purposes: First, I characterize the population composition of Poland from the time of the imperial partitions to the period of communism. I do so for the reason that studying the evolution of Poland's human capital within the framework of the Polish partitions is negligent without paying attention to the simultaneous evolution of the population that inhabits the partitions at various points in time. Skilled migration, for example, can have long-term effects (Hornung, 2014); therefore it could obscure or inflate any partition effects on human capital if it occurred only on one side of a partition border. On the top of that, the question of potentially selective migrations from one partition into another is particularly relevant for the spatial RDD outlined in Section 1.4.1. In short, the Polish population of the partitions has been spatially persistent, making it valid for comparisons along the partition borders as history evolves. However, it has become more homogeneous over time by the expulsion and extinction of ethnic minorities. Second, I sketch the various educational policies that were in place over the course of time.

1.2.1 The imperial period (1795-1918)

1.2.1.1 Population composition

The absence of a Polish state between 1795 and 1918 makes it necessary to rely on information other than nationality in order to characterize the population of the three partitions. The literature (both historic and contemporary) usually infers the number of Poles from either the mother tongue (Polish) or the religion (Roman-Catholic) of the imperial populations. Based on this approach, a compilation of imperial census statistics by Gawryszewski (2005, p.245) suggests that the population of the partition territories that are still part of Poland today was in majority Polish at the turn of the 20th century, with the exception of counties that bordered the remainder of the respective empires where the population was more evenly mixed. The urban areas of the Prussian partition posed another exception; Poznan was the only major city there with a Polish population majority over the German inhabitants. In the Kingdom of Poland, the second-largest population group consisted of Jews, while both Jews and Ukrainians constituted sizable minorities in the part of Galicia that would remain with Poland after WWII.

Despite the pronounced differences between the three partitions, population movements across the partition borders appear to have been a very limited phenomenon. While migration statistics are generally scarce in historical contexts, some publications at least decompose a territory's population with respect to its various places of birth. For example, according to the Russian Imperial Census of 1897, only 1.3% of the inhabitants of the Kingdom of Poland were born in a foreign country (Gawryszewski, 2005). Similarly, only 1.2% of the inhabitants of Galicia were foreign-born according to the population census of the Austrian-Hungarian Empire of 1910 (Bureau der K.K. Statistischen Zentralkommission, 1914). This may not surprise given that the partition boundaries represented actual national borders, which typically inhibit mobility much more than e.g. administrative boundaries within the same country. In addition, Davies (2005) suggests that border enforcement was particular harsh on the Russian side.

1.2.1.2 Education during the imperial period

Both similarities and differences between the three imperial educational systems are reviewed comprehensively by Bukowski (2016). In sum, the Prussian and the Austrian systems were very similar to each other in terms of their local institutions, duration of schooling, and curricula. In addition, school attendance was obligatory for the duration of eight and seven years in the Prussian and Austrian Empire respectively. However, these two empires differed sharply in terms of their respective intentions behind providing education to their Polish citizens: While the Prussian system sought to Germanize the Poles by pushing back the usage of the Polish language in schools, the Austrian-Hungarian Empire granted its Polish citizens the right to operate the schools in Galicia in Polish language. As a consequence, while Prussia enforced high levels of enrollment and accomplished similarly high levels of literacy (as measured in proficiency of the German language) among its population, the schools in the Prussian partition were perceived as a means of German oppression and forced assimilation by the Poles, resulting in several protests and clashes at the beginning of the 20th century, when German nationalist policies intensified. Cinnirella & Schüler (2017) document that within Prussia, counties with a high degree of linguistic polarization were disadvantaged in terms of allocations of public funds for education. In contrast, the Galician schools were perceived as a means of preserving and fostering Polish culture and identity; thus the Poles developed a positive association to them.

The Russian Empire, in turn, combined a sparse provision of educational infrastructure with a hostile attitude of the educational system towards the Polish citizens. Russia, in contrast to the German and Austrian empires, did not introduce compulsory schooling prior to WWI; correspondingly it also provided few educational facilities. Further, the primary schools in the Russian partition typically comprised only three consecutive classes, compared to the seven and eight classes in Austrian and German primary schools. Education was carried out in Russian language only; thereby creating a similar association of education with oppression and forced assimilation on the side of the Poles as in the German partition.

Some aggregate figures give a descriptive idea of the differences in education across the imperial partitions: The population of the Kingdom of Poland amounted to 12.5 million people in 1910. In the same year, the Kingdom of Galicia and Lodomeria had 8 million inhabitants, while the Grand Duchy of Poznan and the part of West Prussia that was to become part of Poland after WWI counted 3.1 million citizens. However, the Kingdom of Poland provided only 3,352 primary schools for its population by contrast with Galicia's 5,580 and Prussia's 4,450. This vast difference in school supply is reflected in the share of primary school students among the total population; this share was only 2.3% in the Kingdom of Poland, while it was 13.5% in Galicia and 19.2% in Prussia. Illiteracy had essentially been eradicated everywhere in the German Empire at the turn of the 20th century when the share of the illiterate population still averaged 59% in the Polish territories of the Russian Empire in 1897 and 56.6% in Galicia in 1900.[2]

1.2.2 The interwar period (1918-1939)

1.2.2.1 Population composition

The Second Republic of Poland, at its core the composite of the three former partition territories, could be equally characterized as a multi-ethnic, multi-lingual, and multi-religious country. This applied in particular to the eastern territories that Poland had annexed after the Polish-Soviet war (1919-1921), as these so-called borderlands (Polish: *Kresy*) housed large groups of Belarusians, Ukrainians, Lithuanians, and Jews. However, these territories did not remain part of Poland after WWII; therefore I exclude them from the following analysis.

The lands of the former Prussian provinces of Posen and West Prussia continued to house

[2]The high share of illiterates in Galicia was driven in particular by its eastern territories. It averaged rather between 30 and 40 % in the western, predominantly Polish-populated counties.

a German minority of considerable size after the Polish takeover. This was not the case in Galicia, where the Austrian-German presence had historically been much lower. More importantly, the size of the German population group in Poland's west did not stay constant after WWI. The emigration movement that took hold of the Germans in the former Prussian provinces soon after the latter's transfer to Poland is the object of study of Blanke (1993). While the Polish constitution of 1921 granted every citizen the right of preserving his nationality and developing his mother-tongue and national characteristics, the auspicious constitutional provisions quickly became subject to interpretations regarding whether they applied to entire minority groups, or just to individuals. For example, German was never recognized as a second official language in the interwar republic. This created the necessity for German officials and associations to demonstrate their proficiency in Polish in order to keep their positions and accreditations; a necessity that many of them could not meet. Unrest among the German population therefore soon resulted in emigration. How much of the German exodus from the former Prussian partition was the result of exaggerated panic among the Germans compared to deliberate emigration for economic reasons or forceful displacement remains a matter of historical debate. Historians also attach a varying degree of reasonable doubt to the unbiasedness of both Prussian and Polish ethnic population statistics. But that emigration took place at a large scale is an undisputed fact (Blanke, 1993): The pre-WWI German population of the Prussian partition and Upper Silesia amounted to approximately 1.1 million. By the end of 1921, about 50% of it had already left Poland. Emigration continued during the subsequent decade in significant numbers: As an orientation, the Polish population census of 1931 reports only about 300,000 native German speakers (and about the same number of Protestants) as remaining in the provinces of Poznan and Pomerania.

These population changes are reflected in the Polish population census of 1921, which decomposes each county's population in terms of its place of birth relative to its current place of residence, similar to the previously cited Russian and Austrian population statistics. Using this information, I calculate the share of each county's population in 1921 that was born in the same partition territory. This share averages 97% in the counties of the former Austrian and Russian partitions, suggesting that very little population relocation took place after the imperial borders had been dissolved. However, only 85% of the inhabitants of the former Prussian partition were born on this partition's territory. In turn, 8% were born outside of Poland's borders as of 1921. While the census does not specify the country of birth, it appears reasonable that many of these foreign-born Poles have immigrated into the former Prussian partition from the neighboring territories that still remained part of the German Reich after the Treaty of Versailles. The former Prussian partition was particularly attracting not only because of its geographical proximity, but also because of the emigration of the German minority, which had created empty space in both rural and urban areas.

1.2.2.2 Education during the interwar period

The first priority of the Second Republic with regard to education was to establish a network of primary (or 'common', Polish: *powszechny*) education in the former Russian partition where obligatory schooling had not previously existed. Consequently, efforts and resources were concentrated on this purpose. Schooling statistics (MWRiOP, 1927) indeed suggest considerable and rapid improvements in primary schooling in the central provinces that are congruent with the former Kingdom of Poland. For example, the number of pupils in primary school rose from about 370,000 in the school year 1910/11 to about 1,200,000 in 1921/22; thereby more than doubling the gross primary enrollment rate from 19.4% to 48.5% over the same time span.

In terms of aligning the different school systems, the number of years of obligatory schooling was set to seven nationwide, thereby decreasing it by one year in the former Prussian partition. Pupils were supposed to start school at the age of seven and hence to complete their primary education at 13. (Krzesniak-Firlej et al., 2014)

While the schools in the former Austrian partition had already been run by Poles in place where Poles constituted a significant population group, the integration of the educational system in the former Prussian partition posed more of a challenge. This has to be understood in the contexts of the pre-WWI Prussian educational policy and the post-WWI emigration process previously described: The educational system in the Prussian partition had been mostly kept in the hands of Germans for fear of Polish nationalism and separatism. According to Blanke (1993), the past German policy resulted in strong drawbacks for the educational system after the Polish annexation. The Polish state, inclined to taking an equally uncompromising stance on the education of minorities as the preceding Prussian administration, saw no reason to pay for German schoolteachers who had previously been employed by the Prussian state. In addition, German teachers also saw a sharp depreciation of their human capital because schools now used Polish as the primary language of instruction. Germans living in the Prussian partition had had no incentive to learn Polish prior to 1918, while Poles had already grown up bilingually, so the former faced disparately higher obstacles to their integration into the Polish state that had decided to treat the German language as disdainful as the Prussian government had treated the Polish language. Consequently, the propensity to emigrate was particularly high among German schoolteachers: Out of 9,000 residing in the former partition in 1918, 8,000 left over the course of the following years. This implies that the school system in Poznania and Pomerania was deprived of a large share of its teaching staff after WWI. It

became necessary to replenish it with Polish teachers first before operations could fully resume.
In order to avoid assimilation, the remaining Germans turned to minority schools, to which they were legally entitled. However, a minimum of forty school children within any school district was necessary for maintaining a school; a requirement that became more and more difficult to meet by the shrinking German minority. Together with the closure of all German teacher-training facilities and administrative obstacles, this led to a decline in the number of German-language primary schools in the former Prussian partition from about 1,250 in the school year 1921/22 to about 480 in 1925/26. Hence, a lack of integration of the remaining German minority into the reshuffled Polish school system might have depressed enrollment in western Poland at least during the first years after WWI.

1.2.3 The communist period (1944-1960)

1.2.3.1 Population composition

Jumping from the interwar period of independence to the era of communism in Poland requires pointing to the drastic population changes and losses that Poland experienced during and after WWII. About 5.2-5.3 million ethnic Poles and Jews are estimated to have perished between 1939 and 1945 (Eberhardt, 2011), amounting to about 15% of Poland's pre-WWII population. Further, several million Poles were deported from territories that the German Empire or the Soviet Union had annexed. They were sent into the German-controlled Generalgouvernement, abducted to Germany for forced labor, or kept in remote areas of the Soviet Union.
However, the liberation of the Polish territory from German occupation did not yet end the mass movement of people all across the Polish lands. In the Potsdam Agreement, the victorious Allied powers decided to move Poland's border westwards. Poland was to cede its eastern territories to the Soviet Union, where they became part of the Lithuanian, Byelorussian, and Ukrainian Soviet Socialist Republics. As compensation, Poland received the remaining German territories of the Prussian state that were located east of the Oder-Neiße line. Due to their geographic location in the Polish People's Republic, these territories have been called the Northern and Western Territories (Polish: *Ziemie Zachodnie i Pólnocne*). In the process of annexing them, their German population was to a large extent forcefully displaced to what was to become East and West Germany respectively. It was replaced by a mix of migrants from central Poland and a group of forcibly resettled Poles from the ceded eastern territories. However, while the complex population structure of the Northern and Western Territories is a promising topic of future research, I exclude them from my analysis, as their boundaries do not intersect with the partition

borders.
The population in the three former partitions, which now formed the territorial core of Poland, was apparently much less affected by these drastic migrations, as census data from 1950 shows that, at least at the level of provinces, on average more than 90% of the inhabitants of the former partitions had already lived in the same respective province in 1939 (GUS, 1955). Thus, despite the numerous population transfers during wartime, large-scale population *replacement* was confined to the Northern and Western Territories.
The Holocaust and the expulsion of Germans raise the question to what extent these events changed the ethnic and religious composition of Poland's population. The communist period makes it difficult to answer this question because of the official concept of a Polish nation united under socialism. However, it is historically accepted that the population of Poland has been much more ethnically homogeneous since the completion of the major population movements after WWII. For example, Eberhardt (2011) estimates that the share of ethnic Poles within the same pre- and post-WWII territories of Poland already rose from 63.9% in 1939 to 85.7% in 1946. It further rose to 97.8% in 1950 due to the continued expulsion of Germans and the emigration of Jewish Holocaust survivors to Israel.

1.2.3.2 Education during the communist period

Needless to say, the loss of lives during WWII also affected Poland's stock of human capital; in particular because both Soviet and German occupants specifically targeted the Polish intelligentsia. Eberhardt (2011) cites evidence that about every third Pole with a university education perished during the war. Educational instructors were decimated with a similar bias towards the highly-educated ones: While about 28.5% of the university lecturers died, 'only' 5.1% of the primary school teachers perished. Educational infrastructure was not spared the intense destruction of physical assets during wartime: In the school year 1944/45, the number of public primary schools amounted to only 86.6% of the number of prewar schools within the same territory of Poland. The fall in the number of schools was roughly equally distributed across the country; only the province of Pomerania operated less than 70% of its facilities in 1944/45 compared to 1937/38. (MWRiOP, 1946)
In addition, the prewar efforts to overcome illiteracy and to provide universal access to education were set back by the turmoil of war and occupation, which had severely constrained any organized education for half a decade. This resulted in 1.4 million illiterates among the Polish population in 1949 as cited by Dobosiewicz (1970). As a consequence, public literacy programs for adults were provided over the course of the next decade.
Not surprisingly, educational policies of the People's Republic were soon directed against

private and religious schools; with the effect of eliminating the latter from the educational sector over the course of the 1950s. After some regulatory turmoil in the early postwar years, the duration of obligatory primary education was set to the prewar level of seven years in 1949. It was extended by one year not until 1961. Educational institutions and curricula were sweepingly harmonized under the socialist command. (Dobosiewicz, 1970)

1.3 Data

1.3.1 Data prior to 1918

The existing literature has to the best of my knowledge not used historical data that predates the partition borders in an attempt to test for the absence of pretreatment discontinuities. In order to perform a rudimentary test, I rely on population data that have been collected during the short-lived reign of the Duchy of Warsaw in 1810. The Duchy had been constituted by Napoleon I in 1807 and comprised, in addition to the core territories around Warsaw, most of what became the Prussian partition in 1815 and some parts of the future Austrian partition. While the process of the Partitions of Poland had already begun in 1772, the partition borders were finalized only in 1815. They arguably did not change considerably along the Austrian territory of Galicia, but they bore only little resemblance to the Prussian-Russian borders drawn before the Napoleonic campaigns. The population data are disaggregated into larger cities, as well as into small settlements. They have been compiled and published by Grossman (1925), who also supplemented them with comparable data on towns and cities in Galicia. Because I do not have a map of the lower-level administrative divisions of the Duchy of Warsaw, I georeference each observation individually and calculate the logarithm of its population as a crude measure of the local prosperity.

Further, I assemble data on the state of public primary education in each of the three imperial partitions of Poland before the outbreak of WWI from Prussian and Russian school censuses and Polish statistics on Galicia (Königlich Preussisches Statistisches Landesamt in Berlin, 1912; Pokrowskoho, 1914; Pilata, 1913). Both the two censuses and the Galician statistics refer to the school year 1911/12. However, not all available information is fully comparable across empires. At the county level, the number of primary schools, primary school teachers, primary school students, and primary school classes are least likely to be affected by varying imperial measurements and definitions. I therefore select these variables for the construction of the imperial-era dataset. While private schools were quite numerous and well-attended in particular during the imperial era, their presence and importance were to an overwhelmingly degree confined to the major cities. Given that not all of my sources report the number of private schools at the county level, I rely on the binary indicator for cities in my regression models to capture any potential differential effects of private schooling.

Given that the Prussian and Austrian school statistics do not specify the age of the pupils enrolled in primary school, I rely on gross primary enrollment rates, defined as the share of primary school students among the school-age population, in the following. However, their calculation is complicated by the fact that only the Prussian school census directly provides the number of children at primary school age along with the number of primary school students. Therefore, I complement the data on the number of students in the Austrian and Russian partitions with population statistics. The Austrian-Hungarian Empire conducted a population census in 1910 that provides the corresponding data on the number of children at school age at the county level (Bureau der K.K. Statistischen Zentralkommission, 1914). The Russian Empire, however, conducted a population census only in 1897 (in fact, the first and last census of the Russian empire). While I obtain figures on the county-level population in the Russian partition in 1910 from Polish sources, these figures are not decomposed by age. Falski (1925) provides the county-level share of children at school age among the population in 1897, with school age defined in terms of the laws of interwar Poland, i.e. seven to 13 years. Applying this definition to the Russian partition prior to 1918 is in a sense arbitrary, as there was no legal school age in the Russian Empire. However, from the age statistics in the Russian school census, I calculate that about 95% of the students in primary school in 1911 fell into the age range 7-13. Therefore, assuming that the population share computed by Falski (1925) did not change considerably between 1897 and 1910, I multiply it with the total county population in 1910 to obtain the number of children at school age in 1910 and thereupon the enrollment rate within the Russian partition.

1.3.2 Data on the interwar period

The importance that the Second Polish Republic attributed to education is reflected in both the amount and the depth of data on schooling and educational attainment that has been collected by Polish statistical agencies at that time.

The first complete and disaggregated picture of primary education in independent Poland emerges from a series of publications by the Central Statistical Office (GUS, 1922) that provide data on the school year 1920/21. This series is followed by a census of primary schools collected in the school year 1925/26 (MWRiOP, 1927). Finally, the Central Statistical Office provides an annual series of school statistics starting in the school year 1932/33 (GUS, 1934).

Poland conducted population censuses in 1921 (GUS, 1927) and 1931 (GUS, 1938) respectively. The first one contains information on literacy, as well as several classifications of educational attainment. The second one is less detailed, but continues the data series on literacy. Consequently, I compute the share of the literate population above primary-school age in both time periods.

The province of Silesia, which consisted of the eastern part of the Prussian province of Upper Silesia, is missing in the census of 1921 because the status of this territory had not been ultimately determined by the time the census was conducted. After its inclusion into Poland, the province further saw numerous changes of administrative boundaries due to its urban character and small territorial units. I therefore exclude Silesia and thereby the only immediate, but relatively short Prussian-Austrian partition border segment from my sample in all time periods.

Similar to before WWI, most of the statistical publications report only the number of students enrolled in primary school in a given year, without providing information on their age. Consequently, I continue to use gross enrollment rates which I calculate for 1921 and 1931 using the population census information on the number of children at school age within each county. For the school year 1925/26, I compute the average school-age population between 1921 and 1931 using the two census datasets in order to obtain an approximate enrollment rate in 1925/26.

1.3.3 Data on the communist period

Population statistics on educational attainment and literacy are taken from the Polish population census of 1960, same as the number of citizens at primary school age (GUS, 1965). I match these with statistics on the number of primary schools and students of the school year 1960/61 (GUS, 1962). I rely only on this time period because of the extensive educational information of the census, the reform of the school system in 1961 and substantial changes in the system of administrative boundaries soon afterwards.

1.3.4 Georeferenced Polish counties

The spatial RDD necessitates a measure of a county's distance to the former imperial borders. The Euclidean distance, i.e. the shortest line connecting two points if there were no obstacles, of a county's centroid to the respective border is a natural candidate.

In order to perform the corresponding calculations in ArcGIS, I mainly rely on a map of the Second Republic of Poland (WIG 1934)[3] for the pre-WWII internal boundaries. The borders between regions in the west and the south of the Second Republic coincide with the former partition borders, such that the latter can be easily reconstructed. While

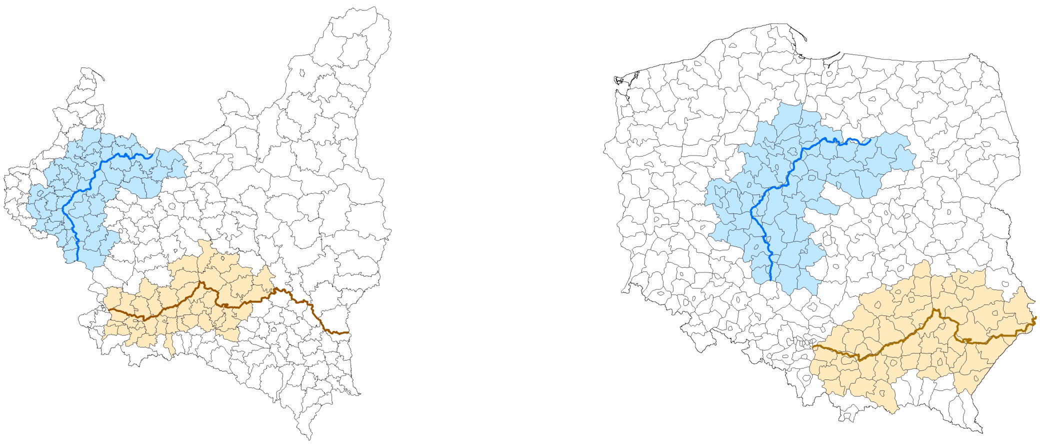
[3]The map was georeferenced and publicly provided by Paul Dziemiela under the Open Data Commons Public Domain Dedication and License. Source: http://dziemiela.com/personal/polish_interwar_geospatial_datab.htm. Accessed on 29 June 2017.

Poland, on the verge of its regained independence, adopted most of its internal administrative boundaries from the three empires that had previously demarcated the Polish territories, several counties have been merged, split up, or rearranged in the course of the decade from 1921 to 1931. I harmonize the boundaries for the sake of using the same number of counties for each time period by drawing on an online repository of legal acts, including administrative changes, of the Polish parliament (ISAP, 2015) and additional georeferenced maps provided by the Mosaic project (MPIDR and CGG, 2011, 2012a, 2012b).

The county boundaries in 1960, likewise obtained from MPIDR and CGG (2012a) do not bear close resemblance to the pre-WWII boundaries anymore. The former partition borders now cut through a small number of counties, which I therefore exclude from the sample. Keeping the bandwidth constant at 65 km, the sample size is reduced by only one county at the Prussian-Russian border in 1960 compared to the pre-WWII sample. However, it increases by 16 at the Austrian-Russian border due to the creation of new counties. While it would be possible to merge some of the new counties in order to bring the sample size closer to the one available for the earlier years, a larger number of observations in 1960 might actually reduce the risk of incorrectly accepting the null hypothesis of a faded partition effect due to imprecise estimation, which is why I leave the county boundaries unchanged.

The counties along the former partition borders within the changing national borders of Poland that are included in my sample as a result of the bandwidth choice of 65 km are displayed in Figure 1.2.

Figure 1.2: Sample counties within national borders of Poland



(a) Partition borders and sample counties within interwar Poland

(b) Partition borders and sample counties within Poland since 1944

1.4 Identification strategy

1.4.1 Spatial regression discontinuity design (RDD)

The Partitions of Poland provide a promising setting for a spatial RDD that allows estimating the causal effect of the partitions on human capital in Poland.
The basic idea behind a RDD is that if individual assignment to treatment is determined by an assignment variable exceeding a threshold and if individuals have imprecise control over the assignment variable, then assignment to treatment is randomized for individuals just below and above the threshold. A well-known example from the literature on the economics of education is that if grant eligibility is tied to student performance in a test, then students are unlikely to have precise control over their test scores such that their exact test score is random within a certain neighborhood. This implies that students who score marginally above or below the grant threshold are randomly assigned to treatment and control. Moreover, they are likely to be similar to each other in terms of both observable and unobservable pretreatment characteristics, rendering them valid comparison groups. Indeed, if variation in treatment status is randomized around the threshold, then all characteristics determined prior to the realization of the assignment variable should evolve smoothly around the threshold. (Lee & Lemieux, 2010)
In the spatial context, the assignment variable is typically understood as the distance of a spatial object to a multidimensional discontinuity in space such as a border. The assignment variable exceeding the threshold then translates into crossing this border from one territory into another, with the respective territorial affiliation corresponding to either treatment or control status. Consequently, the causal treatment effect can be identified by comparing observations on both sides of the border, but close to it. However, this requires that the units of observation, for example households or firms, could neither deliberately manipulate the course of the border, nor change their location as response to the border in order to receive (or avoid) treatment. Further, it implies that the border had to be constructed exogenously with regard to the spatial distribution of predetermined variables that influence the outcome of interest. Consequently, these variables should evolve smoothly around the border (Dell, 2010).
In the geographical-historical context of Poland, the treatment variable is the identity of the respective partition power. Given that no territory under Polish self-government existed during the time of the partitions, I consider the Russian partition in central Poland as the control group in the following. The assignment variable is therefore the distance to the two borders separating the Russian and Austrian partitions and the Russian and Prussian partitions respectively, with treatment being assigned by crossing from the Russian into either the Austrian or Prussian partition. In order to identify the causal effect

of the partition borders on education, these borders had to be drawn by the partitioning powers in disregard of local conditions that would influence education. These conditions, whether observable or unobservable, should therefore change smoothly at the borders under investigation. Migrations across partitions, in particular if they occurred due to more promising educational opportunities in the Austrian and Prussian empires, would represent a form of manipulating assignment to treatment.

1.4.2 Econometric specification

The literature that employs spatial RDDs distinguishes a one-dimensional and a two-dimensional parametric approach. In the one-dimensional approach, the forcing variable, in my case an observation's distance to a partition border, enters the regression model linearly. Interacting the distance measure $Distance_i$ with an $Empire_i$ dummy indicating the treated partition territory the county is located in (i.e. either Austria or Prussia) further allows the effect of the distance to vary at each side of the border. Adding the county's longitude X_i, latitude Y_i, and a vector of controls $\mathbf{C_i}$ results in the following regression model that can be estimated by OLS:

$$y_{it} = \alpha Empire_i + \beta_1 Distance_i + \beta_2 Empire_i \cdot Distance_i + \gamma_1 X_i + \gamma_2 Y_i + \delta \mathbf{C_i} + \epsilon_{it} \tag{1.1}$$

The parameter α then identifies the causal effect of either the Austrian or the Prussian empire on the outcome y in period t, depending on the border at which the parameter is estimated. The two-dimensional approach proposed by Dell (2010) is not interested in the direct effect of distance to the border as the forcing variable. Instead, it uses a polynomial of latitude and longitude $f(X_i, Y_i)$ in order to flexibly control for a county's geographic location along the border:

$$y_{it} = \alpha Empire_i + f(X_i, Y_i) + \delta \mathbf{C_i} + \epsilon_{it} \tag{1.2}$$

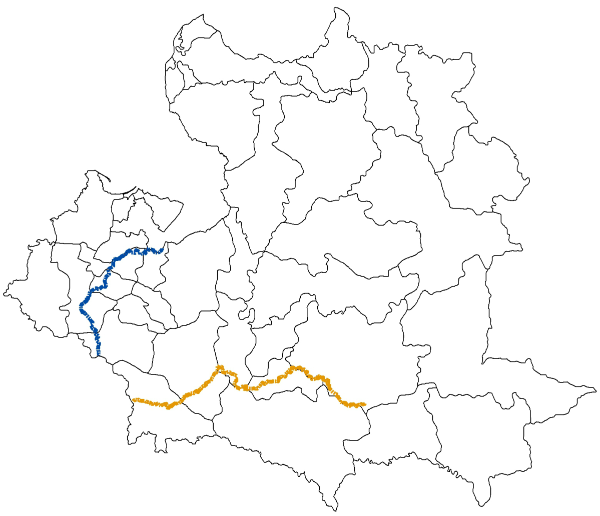
Given that I rely on county-level data in the following, the number of observations within a reasonable distance to the partition borders is relatively small. It precludes the utilization of more data-intensive nonparametric methods for estimating the effect of spatial discontinuities.

The choice of the bandwidth in both specifications involves a trade-off given that a wider bandwidth increases the number of observations and thereby statistical power, but also casts more doubt on the linearity assumption on the forcing variable or the ability of the polynomial to appropriately control for the geographic characteristics. For my baseline specification, I choose a bandwidth of 65 km on each side of the partition border under consideration. Grosfeld & Zhuravskaya (2015) and Bukowski (2016) choose narrower bandwidths of 60 and 50 km respectively; however, their data are disaggregated to the municipality level, thereby providing more observations. Similarly, the choice of the functional form of the longitude-latitude-polynomial $f(X_i, Y_i)$ in the two-dimensional specification invokes a trade-off: Raising the order of the polynomial increases flexibility, but it also amplifies the threat of overfitting the data. Following the recommendation of Gelman & Imbens (2016), I employ a (relatively low-order) quadratic polynomial of latitude and longitude for Equation 1.2.

1.4.3 Validity of the spatial RDD

In order to provide a simple visual impression, I overlay the internal divisions of the Polish state prior to its partitions with the final partition borders of 1815. Figure 1.3 shows that the regional borders of the Polish-Lithuanian Commonwealth in 1770 (MPIDR and CGG, 2012a) and the partition borders are hardly congruent and in areas where they do overlap, they mostly follow rivers. Further, there is no historic evidence that the local

Figure 1.3: Borders of the Polish-Lithuanian Commonwealth and partition borders



Polish population along the partition borders had any possibility for manipulating their assignment to treatment, i.e. for influencing the decision on which side of an imperial border their municipality or city would be located after 1815. Indeed, the absolutist character of the three empires at the time when they agreed upon the partitions makes it unlikely that their subjects were granted any say in these decisions.

Next, I present evidence on potential pretreatment discontinuities in terms of the partitions' (log) population in 1810. Results from estimating the one-dimensional RDD do not indicate the existence of such discontinuities (Table 1.1). While there is a statistically significant estimate at the Austrian border (column 5) when all observations on both sides of the border are included, the significance vanishes as soon as bandwidths narrower than 100km on each side of the border are chosen. Switching to the two-dimensional specification (Table 1.2) does not yield any significant estimate.

Table 1.1: Log population at partition borders in 1810 (One-dimensional RDD)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable	Log Population in 1810							
Prussian Side = 1	0.036	-0.126	0.078	-0.017				
	(0.137)	(0.184)	(0.234)	(0.267)				
Austrian Side = 1					0.323**	0.247	0.131	0.077
					(0.141)	(0.157)	(0.184)	(0.219)
Observations	621	245	165	137	621	251	187	145
R-squared	0.140	0.125	0.136	0.076	0.152	0.163	0.198	0.227
Distance, Distance*Partition	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude/Longitude, City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	.	100km	65km	50km	.	100km	65km	50km

Notes: One-dimensional RDD. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.2: Log population at partition borders in 1810 (Two-dimensional RDD)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable	Log Population in 1810							
Prussian Side = 1	-0.088	-0.120	-0.002	-0.064				
	(0.148)	(0.179)	(0.220)	(0.237)				
Austrian Side = 1					0.119	0.251	0.195	0.263
					(0.141)	(0.155)	(0.184)	(0.204)
Observations	538	245	165	137	470	251	187	145
R-squared	0.128	0.136	0.147	0.077	0.183	0.179	0.217	0.243
2nd order Polynomial, City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	.	100km	65km	50km	.	100km	65km	50km

Notes: Two-dimensional RDD. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

While I am aware of the simplicity of my pretreatment measure, it gives no evident reason to question the smoothness of the population distribution at the designated partition borders before 1815. In addition, the claim of the exogeneity of the partition borders is supported by Grosfeld & Zhuravskaya (2015), who review numerous historical sources that suggest that the partition borders did not reflect preexisting economic, ethnic or religious divisions.

While the pretreatment distribution of variables such as education or income cannot be tested, it is common practice to check whether geographical characteristics are smooth at the borders of interest. Discontinuities in geography may indicate that local geography played a role in constructing the border and might therefore confound the outcome. I test for geographical discontinuities by estimating Equation 1.1 and Equation 1.2 with altitude, precipitation, and temperature as the left-hand-side variables respectively. Data for the three variables are obtained from WorldClim 1.4 (Hijmans et al., 2005). Results are reported in Table 1.3. Using a bandwidth of 65 kilometers, I find statistically significant discontinuities in all three geographic variables at the Austrian-Russian partition border in both the one- (Panel A) and the two-dimensional (Panel B) regression design. My estimates are of similar magnitude to those of Bukowski (2016), who refers the discontinuities to the local riverbed of the Vistula. Indeed, when plotting the three variables against the distance to the border, it becomes apparent that altitude and precipitation actually steadily increase on both sides of the border despite the negative sign of the Austrian partition coefficient (Figure A.1 in the appendix to this chapter). This suggests that while counties on the Austrian side of the border are on average more elevated than those on the Russian side, the partition border does not reflect an abrupt change that lifts all observations on the Austrian territory to a different level of altitude. The discontinuities in geography are smaller and mostly insignificant across the Prussian-Russian border (Panel C), except for temperature in the two-dimensional specification (Panel D).

Table 1.3: Discontinuities in geographic characteristics

Dep. Variable	(1) Altitude (m)	(2)	(3) Precipitation (mm)	(4)	(5) Temperature (°C)	(6)
Austrian-Russian Border						
Panel A: One-Dimensional RDD						
Austrian Side = 1	-102.805***	-94.181***	-31.772	-31.391**	0.631***	0.549***
	(34.641)	(29.948)	(34.414)	(13.800)	(0.168)	(0.164)
Observations	44	44	44	44	44	44
R-squared	0.370	0.647	0.417	0.909	0.403	0.492
Distance, Distance*Austrian Side	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: Two-dimensional RDD						
Austrian Side = 1	-117.012***	-110.687***	-48.946***	-46.137***	0.657***	0.631***
	(29.096)	(28.448)	(9.216)	(8.907)	(0.164)	(0.164)
Observations	44	44	44	44	44	44
R-squared	0.741	0.760	0.960	0.965	0.484	0.498
2nd Order Polynomial	Yes	Yes	Yes	Yes	Yes	Yes
Prussian-Russian Border						
Panel C: One-Dimensional RDD						
Prussian Side = 1	11.001	1.221	0.021	-2.993	-0.047	-0.051
	(16.798)	(8.568)	(15.112)	(8.610)	(0.290)	(0.063)
Observations	54	54	54	54	54	54
R-squared	0.213	0.802	0.003	0.651	0.010	0.950
Distance, Distance*Prussian Side	Yes	Yes	Yes	Yes	Yes	Yes
Panel D: Two-dimensional RDD						
Prussian Side = 1	10.694	11.253	-0.737	-0.483	-0.133**	-0.136**
	(9.330)	(8.501)	(6.940)	(6.742)	(0.065)	(0.060)
Observations	54	54	54	54	54	54
R-squared	0.714	0.792	0.664	0.701	0.916	0.933
2nd Order Polynomial	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes

Notes: One- and two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

It appears unlikely that altitude or precipitation have any direct effect on e.g. school construction or education. However, the geographic discontinuities might reflect similar discontinuities in agricultural suitability of the territories in my sample. Agricultural suitability, in turn, might affect the opportunity cost of schooling both before and after the erection of the partition borders. I test for discontinuities in agricultural suitability by proxying the latter first with the Caloric Suitability Index (CSI) by Galor & Özak (2016) and second with the historical croplands dataset by Ramankutty et al.(1999). The CSI provides four grid cell-level estimates of caloric suitability: Average potential caloric yield attainable given the set of crops that are suitable for cultivation pre-/post-1500CE and maximum potential caloric yield attainable given the set of crops that are suitable for cultivation pre-/post-1500CE. I select the two estimates that refer to the post-1500CE era. The historical croplands dataset ignores potential or actual yields and instead provides estimates of permanent cropland areas (as the share of cropland in a grid cell's total land cover) over several centuries, from which I select the years 1800 and 1900. Results are presented in Table 1.4. The estimated discontinuity in both average and potential crop yields are large and significant at the Russian-Austrian border in both RDDs, suggesting a substantially higher yield at the Austrian side (columns 1 and 3 in Panel A and B). However, the effect becomes negative and insignificant when altitude, precipitation, and temperature are included as geographic controls (columns 2 and 4).

Table 1.4: Discontinuities in agriculture

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable	Average Caloric Yield		Optimal Caloric Yield		Cropland 1800		Cropland 1900	
Austrian-Russian Border								
Panel A: One-Dimensional RDD								
Austrian Side = 1	186.077**	-18.295	407.150*	-70.439	-0.017	-0.069	-0.033	-0.101
	(91.586)	(56.696)	(229.682)	(154.235)	(0.047)	(0.045)	(0.060)	(0.060)
Observations	44	44	44	44	44	44	44	44
R-squared	0.405	0.830	0.446	0.790	0.538	0.770	0.500	0.751
Distance, Distance*Austrian Side	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude/Longitude, City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo Controls	No	Yes	No	Yes	No	Yes	No	Yes
Panel B: Two-dimensional RDD								
Austrian Side = 1	199.198***	-21.583	498.364***	60.065	0.015	-0.015	0.014	-0.026
	(61.451)	(21.751)	(127.118)	(81.472)	(0.032)	(0.042)	(0.042)	(0.056)
Observations	44	44	44	44	44	44	44	44
R-squared	0.663	0.950	0.706	0.949	0.818	0.828	0.791	0.804
2nd Order Polynomial, City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo Controls	No	Yes	No	Yes	No	Yes	No	Yes
Prussian-Russian Border								
Panel C: One-Dimensional RDD								
Prussian Side = 1	-80.511*	-90.582**	-227.352*	-214.485**	0.006	0.009	0.007	0.012
	(45.155)	(36.098)	(118.438)	(103.037)	(0.019)	(0.016)	(0.025)	(0.021)
Observations	54	54	54	54	54	54	54	54
R-squared	0.636	0.745	0.650	0.753	0.417	0.566	0.411	0.562
Distance, Distance*Prussian Side	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Latitude/Longitude, City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo Controls	No	Yes	No	Yes	No	Yes	No	Yes
Panel D: Two-dimensional RDD								
Prussian Side = 1	-45.711*	-27.849	-197.538**	-82.676	-0.020	-0.014	-0.027	-0.019
	(23.579)	(20.215)	(81.034)	(61.883)	(0.019)	(0.019)	(0.024)	(0.026)
Observations	54	54	54	54	54	54	54	54
R-squared	0.882	0.918	0.801	0.902	0.529	0.579	0.529	0.576
2nd Order Polynomial, City	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo Controls	No	Yes	No	Yes	No	Yes	No	Yes

Notes: One- and two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

At the Prussian-Russian border, the effect on crop yield is negative and significant in the one-dimensional specification both with and without the geographic controls (Panel C), while the latter render the effect insignificant in the two-dimensional specification (Panel D). The negative effect in the one-dimensional specification is hardly of economic importance: For example, average caloric yield is 2,025 at the Prussian-Russian border;

the estimated discontinuity of 91 amounts to less than five percent of this average. This difference is unlikely to account for the large differences in schooling between the Prussian and the Russian partition. In addition, both employment in agriculture and school enrollment were more prevalent in the Prussian partition than in the Kingdom of Poland during the imperial era, which does not hint at a relevant trade-off between the two. Furthermore, there are no statistically significant discontinuities in historical cropland at any border neither in 1800 nor 1900 (columns 5-8). Nevertheless, I include all three geographic variables as controls in the following regressions for both borders.

The relevance of population movements as a threat to identification has already been discussed in Section 1.2.1.1: There is no evidence for selective migrations (or large migrations of any kind) across the various partitions borders during the imperial era, implying a very limited potential for treatment status manipulation.

1.5 Results

1.5.1 Main results

Table 1.5 presents estimates of the discontinuity in primary enrollment at the Prussian-Russian partition border over the years 1911 to 1961 using the two-dimensional RDD.[4] In addition to the estimates of the partition effect, I also report the mean of the dependent variable at the Russian side of the border for each time period in the sample. Keeping the bandwidth fixed at 65 kilometers, primary enrollment is estimated to be more than 80 pp higher in the Prussian partition in 1911/12 when the empires were still existent. Less than ten years later and two years in the reinstated Polish Republic, the difference is roughly cut in half. Both findings are consistent with the descriptives cited in Section 1.2.1.2. The partition effect further falls below 10 pp in 1925/26 and loses significance, while it shows a slight rebound in 1931/32 before it fades entirely in 1960/61. The various controls for longitude/latitude, cities and geography increase the precision of the estimates, but they do not impact their size. Besides the partition effect, the steadily increasing mean of enrollment in the Russian partition over time further suggests that the dwindling partition effect is indeed the result of increasing enrollment in the Russian partition instead of a potential convergence of both partitions to a rather mediocre level of enrollment: In 1931/32, (gross) primary enrollment averages already close to 100 percent in the Russian partition. I visualize the development of the Prussian partition effect across four of the five time periods in Figure 1.4.

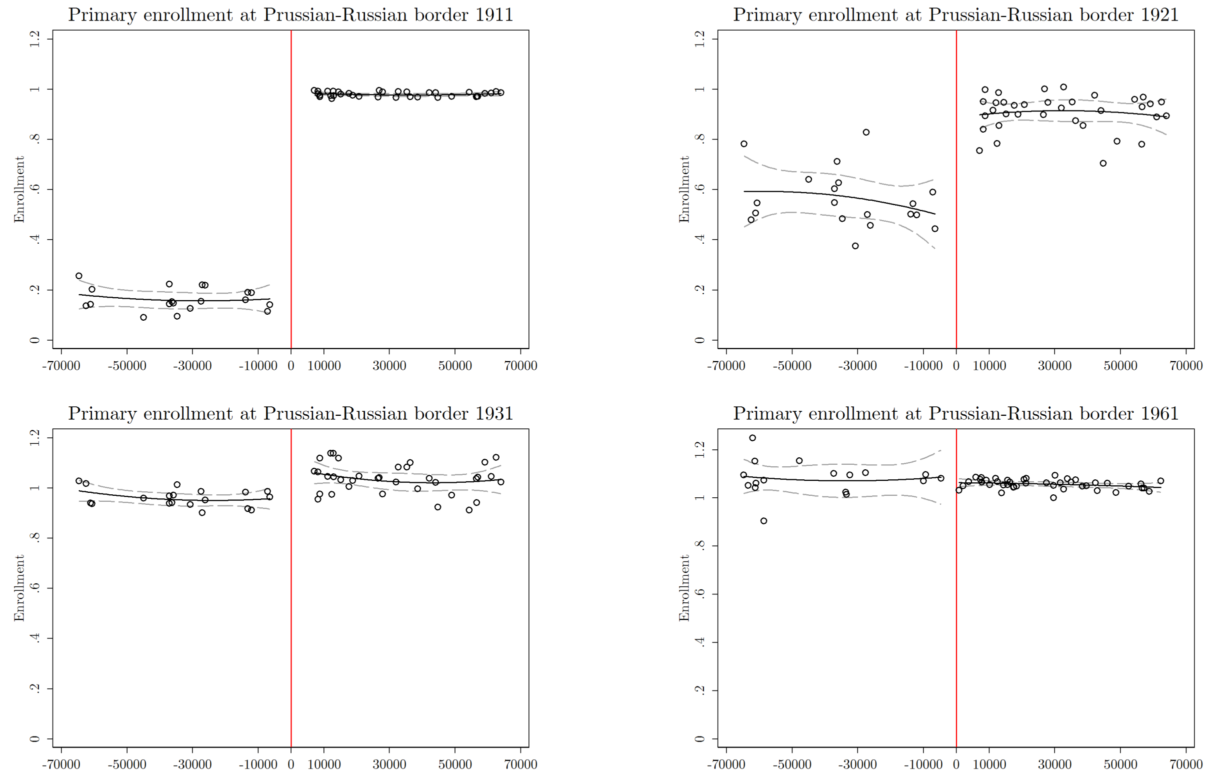
[4]Estimating the one-dimensional specification delivers very similar results and because the distance to the border is not of particular interest in my setting, I delegate all results obtained from the one-dimensional RDD to the appendix to this chapter.

Table 1.5: Primary enrollment at Prussian-Russian border 1911-1961

	(1)	(2)	(3)
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Prussian Side = 1	0.822***	0.822***	0.832***
	(0.016)	(0.016)	(0.015)
Observations	54	54	54
R-squared	0.996	0.996	0.997
Mean on Russian Side	0.164	0.164	0.164
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Prussian Side = 1	0.354***	0.355***	0.378***
	(0.042)	(0.042)	(0.044)
Observations	54	54	54
R-squared	0.875	0.888	0.902
Mean on Russian Side	0.562	0.562	0.562
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Prussian Side = 1	0.060	0.064*	0.066*
	(0.040)	(0.033)	(0.033)
Observations	54	54	54
R-squared	0.185	0.645	0.647
Mean on Russian Side	0.711	0.711	0.711
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Prussian Side = 1	0.097***	0.098***	0.098***
	(0.027)	(0.026)	(0.026)
Observations	54	54	54
R-squared	0.410	0.528	0.550
Mean on Russian Side	0.961	0.961	0.961
Dep. Variable	Primary Enrollment 1961	Primary Enrollment 1961	Primary Enrollment 1961
Prussian Side = 1	-0.027	-0.027	-0.031
	(0.026)	(0.027)	(0.037)
Observations	53	53	53
R-squared	0.183	0.187	0.215
Mean on Russian Side	1.081	1.081	1.081
2nd Order Polynomial	Yes	Yes	Yes
City Dummy	No	Yes	Yes
Geographic Controls	No	No	Yes

Notes: Two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1.4: Discontinuities in primary enrollment at Prussian-Russian border



Y axis: Share of primary enrollment. X axis: Distance to the border in kilometers. Negative distance indicates Russian partition. Bandwidth: 65km.

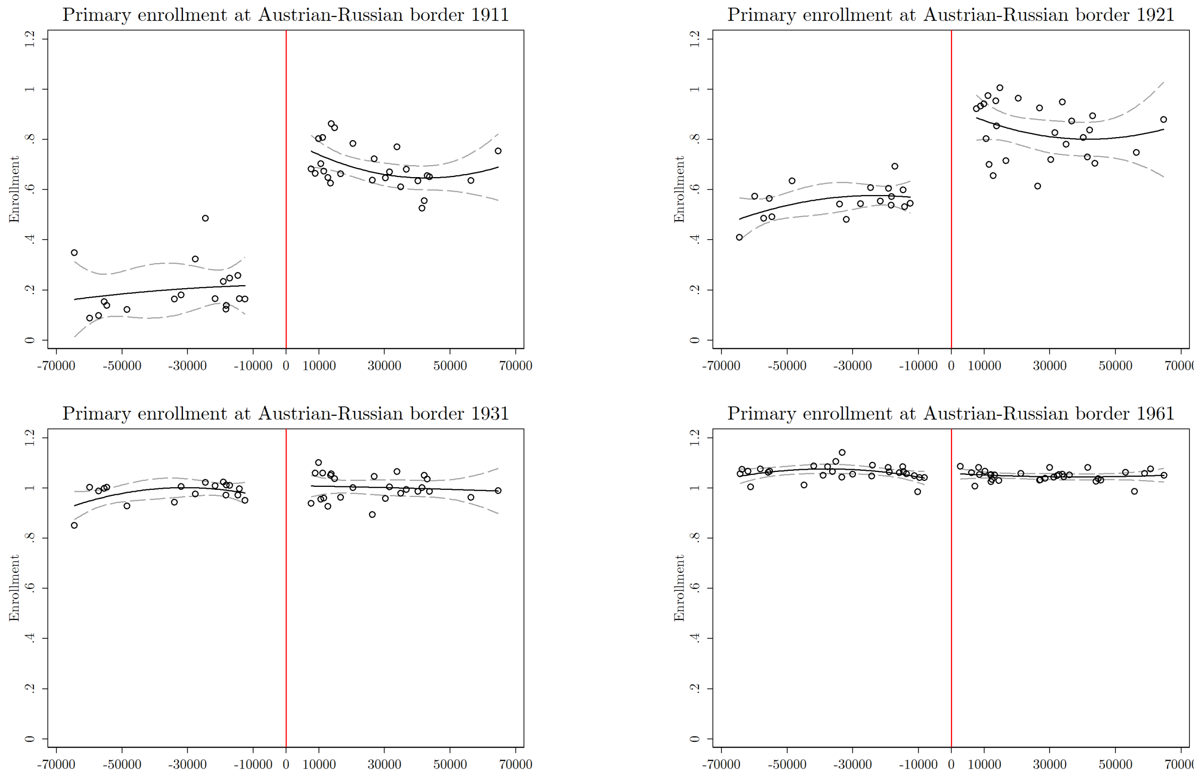
The estimated effect of the discontinuity in enrollment at the Austrian-Russian border (Table 1.6) is of smaller magnitude in 1911/12 in comparison with the Prussian partition effect, reflecting the fact that school attendance in the Austrian-Hungarian Empire was not nearly as universal as in Prussia. Nevertheless, the effect still amounts to about 50 pp. Moreover, the temporal pattern of decreasing relevance of the partitions is strictly monotonic at the Austrian-Russian border: With every subsequent time period, the estimated effect becomes smaller in magnitude; in 1931/32, it is not anymore statistically different from zero and it shows no sign of a resurgence in 1960/61. The controls do not alter this pattern; only the geographic controls render the effect insignificant already in 1925/26. Given that enrollment averages 98% in 1931/32, the counties on both sides of the Austrian-Russian partition border appear to converge to a high level of enrollment. A graphical representation of the results can be found in Figure 1.5.

Table 1.6: Primary enrollment at Austrian-Russian border 1911-1961

	(1)	(2)	(3)
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Austrian Side = 1	0.495***	0.499***	0.400***
	(0.032)	(0.033)	(0.043)
Observations	43	43	43
R-squared	0.926	0.927	0.945
Mean on Russian Side	0.194	0.194	0.194
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Austrian Side = 1	0.226***	0.232***	0.156***
	(0.036)	(0.037)	(0.045)
Observations	44	44	44
R-squared	0.835	0.838	0.870
Mean on Russian Side	0.554	0.554	0.554
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Austrian Side = 1	0.101**	0.100**	0.028
	(0.044)	(0.047)	(0.050)
Observations	44	44	44
R-squared	0.381	0.382	0.428
Mean on Russian Side	0.627	0.627	0.627
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Austrian Side = 1	-0.007	0.005	-0.009
	(0.020)	(0.019)	(0.027)
Observations	44	44	44
R-squared	0.441	0.653	0.662
Mean on Russian Side	0.981	0.981	0.981
Dep. Variable	Primary Enrollment 1961	Primary Enrollment 1961	Primary Enrollment 1961
Austrian Side = 1	-0.012	-0.010	-0.002
	(0.011)	(0.012)	(0.014)
Observations	59	59	59
R-squared	0.114	0.139	0.192
Mean on Russian Side	1.062	1.062	1.062
2nd Order Polynomial	Yes	Yes	Yes
City Dummy	No	Yes	Yes
Geographic Controls	No	No	Yes

Notes: Two-dimensional RDD. 65 km bandwidth. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1.5: Discontinuities in primary enrollment at Austrian-Russian border



Y axis: Share of primary enrollment. X axis: Distance to the border in kilometers. Negative distance indicates Russian partition. Bandwidth: 65km.

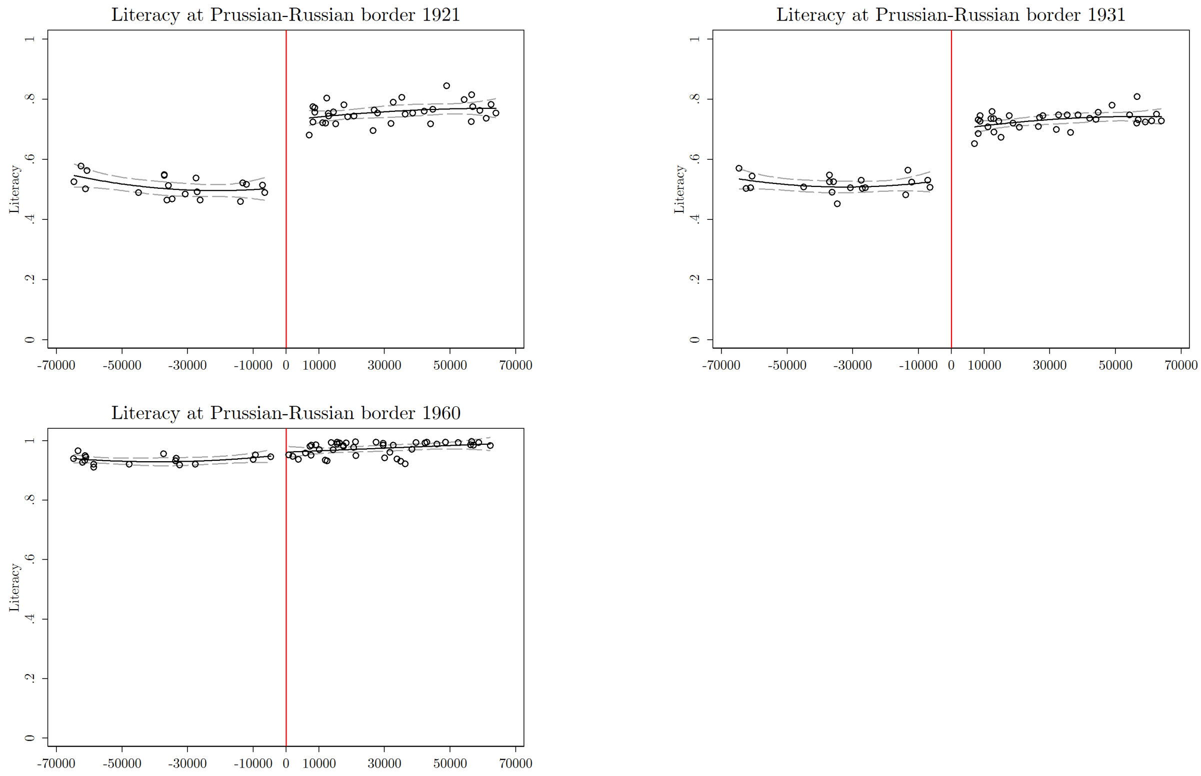
The effect of the partitions on literacy as a measure of accumulated human capital turns out to be more persistent. Table 1.7 shows that literacy in 1921 is 25 pp higher on the Prussian side of the partition border than on the Russian one. Interestingly, literacy among the population of the former Prussian partition is far from universal in 1921, but amounts to only 75%. This may suggest that prior to WWI, universal attendance of a (German-speaking) school and thereby proficiency in the German language did not translate into a comparable spoken and written proficiency of the Polish language, despite private efforts and initiatives of the Polish population. The literacy advantage of the Prussian partition narrows by only 5 pp in 1931. Thirty years later, it is not statistically different from zero anymore. The corresponding graphs of the evolution of the discontinuity in literacy at the Prussian-Russian border are shown in Figure 1.6.

Table 1.7: Literacy at Prussian-Russian border 1921-1960

	(1)	(2)	(3)
Dep. Variable	Share of Literates 1921	Share of Literates 1921	Share of Literates 1921
Prussian Side = 1	0.246***	0.245***	0.244***
	(0.015)	(0.013)	(0.012)
Observations	54	54	54
R-squared	0.942	0.960	0.967
Mean on Russian Side	0.509	0.509	0.509
Dep. Variable	Share of Literates 1931	Share of Literates 1931	Share of Literates 1931
Prussian Side = 1	0.197***	0.196***	0.192***
	(0.012)	(0.010)	(0.011)
Observations	54	54	54
R-squared	0.954	0.973	0.976
Mean on Russian Side	0.517	0.517	0.517
Dep. Variable	Share of Literates 1960	Share of Literates 1960	Share of Literates 1960
Prussian Side = 1	0.003	0.003	0.011
	(0.010)	(0.009)	(0.009)
Observations	53	53	53
R-squared	0.735	0.787	0.862
Mean on Russian Side	0.936	0.936	0.936
2nd Order Polynomial	Yes	Yes	Yes
City Dummy	No	Yes	Yes
Geographic Controls	No	No	Yes

Notes: Two-dimensional RDD. Bandwidth 65. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1.6: Discontinuities in literacy at Prussian-Russian border



Y axis: Share of adult literates. X axis: Distance to the border in kilometers. Negative distance indicates Russian partition. Bandwidth: 65km.

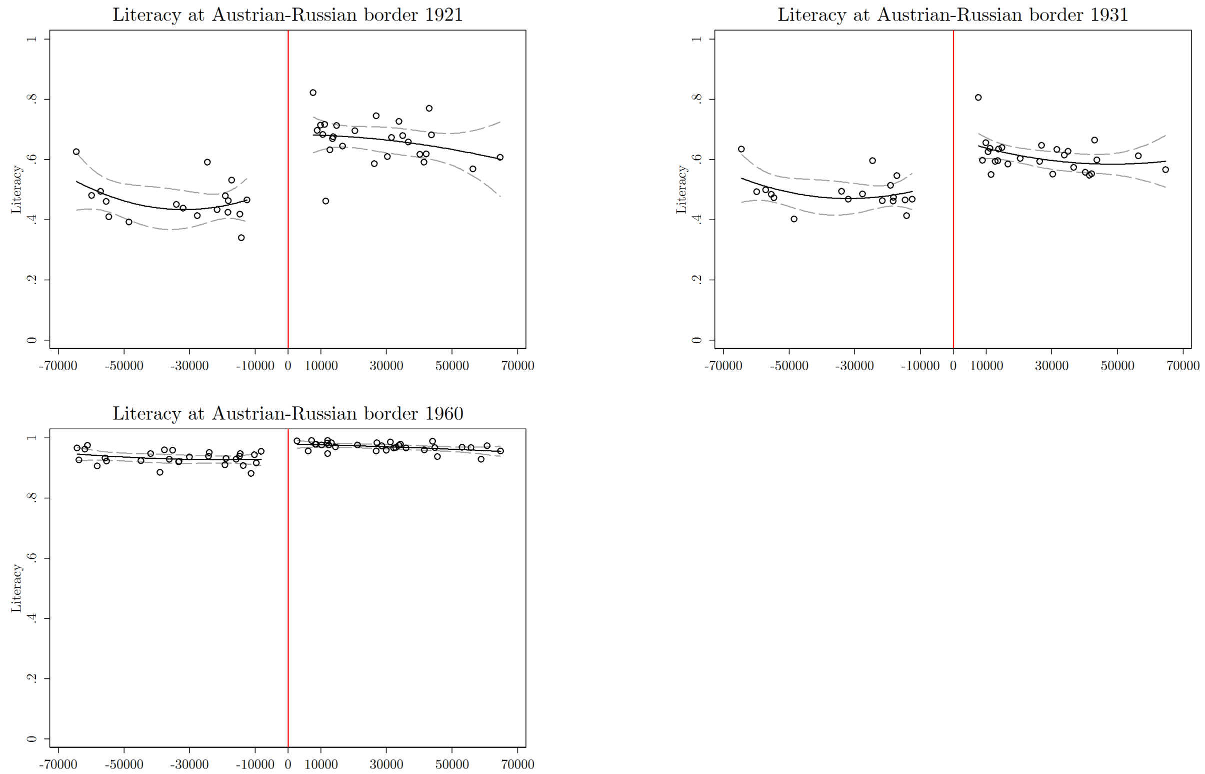
The conjecture that the benefit of high enrollment in the Prussian partition prior to 1918 might have been diminished by the fact that German was the language of instruction is supported by the finding that the partition effect on literacy is of similar magnitude at the Austrian-Russian border as at the Prussian-Russian border in 1921 (Table 1.8). This suggests that the lower enrollment in the Austrian-Hungarian Empire was indeed compensated (in terms of literacy in 1921) by having already used Polish as the language of instruction. The effect of the partition decreases more than at the Prussian-Russian border between 1921 and 1931; however, given the very modest increase in average literacy on the Russian side, it rather seems that literacy actually decreased in the Austrian partition. In 1960, average literacy in the counties on the former Russian side of the Austrian-Russian border is about as high as on the former Russian side of the Prussian-Russian border. The Austrian partition effect is still very pronounced in terms of statistical significance, but with about 5 pp, the remaining difference seems rather small. The corresponding graphical representation at the Austrian-Russian border is contained in Figure 1.7.

Table 1.8: Literacy at Austrian-Russian border 1921-1960

	(1)	(2)	(3)
Dep. Variable	Share of Literates 1921	Share of Literates 1921	Share of Literates 1921
Austrian Side = 1	0.242***	0.228***	0.170***
	(0.027)	(0.026)	(0.042)
Observations	44	44	44
R-squared	0.836	0.886	0.908
Mean on Russian Side	0.462	0.462	0.462
Dep. Variable	Share of Literates 1931	Share of Literates 1931	Share of Literates 1931
Austrian Side = 1	0.156***	0.138***	0.093***
	(0.023)	(0.019)	(0.027)
Observations	44	44	44
R-squared	0.676	0.865	0.899
Mean on Russian Side	0.491	0.491	0.491
Dep. Variable	Share of Literates 1960	Share of Literates 1960	Share of Literates 1960
Austrian Side = 1	0.046***	0.041***	0.040***
	(0.007)	(0.007)	(0.007)
Observations	59	59	59
R-squared	0.574	0.740	0.769
Mean on Russian Side	0.934	0.934	0.934
2nd Order Polynomial	Yes	Yes	Yes
City Dummy	No	Yes	Yes
Geographic Controls	No	No	Yes

Notes: Two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1.7: Discontinuities in literacy at Austrian-Russian border



Y axis: Share of adult literates. X axis: Distance to the border in kilometers. Negative distance indicates Russian partition. Bandwidth: 65km.

1.5.2 Robustness

I present a series of robustness checks to my previous estimates in Table 1.9. Column 1 shows again the results from estimating the two-dimensional RDD with enrollment as the dependent variable. Column 2 shows the effect of adding partition border segment dummies to the regression model. The border segments coincide with province borders before and after WWI, as the latter were constructed along the former partition borders. Along each partition border, there are two provinces on every side of each border. Neither direction, size, nor significance of the estimates are altered by the inclusion of the border segment dummies, with the exception of the school year 1920/21 at the Austrian-Russian border, where the discontinuity becomes insignificant. Finally, I lower the bandwidth to 50 instead of 65 km on each side of every partition border. Doing so reduces the sample size by about 20%. As displayed in column 3, the estimates rather increase at the Austrian-Russian border in 1911/12 and 1920/21, while the significance disappears in 1925/26. At the Prussian-Russian border, they remain the same as for the 65 km bandwidth in terms of size and significance. Taken together, there is no evidence to doubt the robustness of the initially large, but then fading partition effects at both borders.

Table 1.9: Robustness of the two-dimensional RDD

	(1)	(2)	(3)
Prussian-Russian Border			
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Prussian Side = 1	0.832***	0.882***	0.825***
	(0.015)	(0.021)	(0.020)
Observations	54	54	42
R-squared	0.997	0.997	0.997
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Prussian Side = 1	0.378***	0.332***	0.355***
	(0.044)	(0.074)	(0.054)
Observations	54	54	42
R-squared	0.902	0.906	0.912
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Prussian Side = 1	0.066*	0.037	0.056
	(0.033)	(0.053)	(0.041)
Observations	54	54	42
R-squared	0.647	0.713	0.677
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Prussian Side = 1	0.098***	0.086**	0.073**
	(0.026)	(0.037)	(0.033)
Observations	54	54	42
R-squared	0.550	0.552	0.6316
Dep. Variable	Primary Enrollment 1961	Primary Enrollment 1961	Primary Enrollment 1961
Prussian Side = 1	-0.031	0.029	0.011
	(0.037)	(0.057)	(0.025)
Observations	53	53	39
R-squared	0.215	0.269	0.442
Austrian-Russian Border			
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Austrian Side = 1	0.412***	0.365***	0.401***
	(0.043)	(0.077)	(0.048)
Observations	44	44	37
R-squared	0.930	0.933	0.945
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Austrian Side = 1	0.156***	0.102	0.178***
	(0.045)	(0.063)	(0.056)
Observations	44	44	37
R-squared	0.870	0.888	0.860
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Austrian Side = 1	0.028	0.104	0.053
	(0.050)	(0.083)	(0.057)
Observations	44	44	37
R-squared	0.428	0.483	0.515
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Austrian Side = 1	-0.009	-0.001	-0.001
	(0.027)	(0.026)	(0.033)
Observations	44	44	37
R-squared	0.662	0.742	0.618
Dep. Variable	Primary Enrollment 1961	Primary Enrollment 1961	Primary Enrollment 1961
Austrian Side = 1	-0.002	0.013	0.019
	(0.014)	(0.018)	(0.021)
Observations	59	59	47
R-squared	0.192	0.213	0.238
2nd order polynomial	Yes	Yes	Yes
City, Geo Controls	Yes	Yes	Yes
Border Segment Dummies	No	Yes	No
Bandwidth (km)	65	65	50

Notes: Two-dimensional RDD. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

1.6 Mechanisms

1.6.1 Changes in the population composition

I assess the potential importance of the pre-WWII population composition for the evolution of schooling and human capital in two ways: First, I add the shares of Jews and Protestants among a county's population in every time period from 1911 to 1931 to Equation 1.2. Protestantism and ethnic as well as linguistic affinity to Germany were sufficiently closely correlated in the Prussian partition in order for Protestantism to serve as a valid proxy for both of the latter. Second, I control for the share of a county's population in 1921 that was not born in the same former partition that it currently resides in. As indicated in Subsection 1.2.2.1, this share is negligible at the Austrian-Russian border, but not at the Prussian-Russian counterpart due to the partial emigration of the German population after WWI. Assuming that Polish immigrants into the former Prussian partition were attracted to those areas that had previously been abandoned by the Germans, this measure captures rather the effect of recent population replacement, in distinction to the effect of mere declining importance of an ethnic-religious minority picked up by the two religious shares.

In general, the population composition controls do not affect the estimated effects of the Prussian partition sufficiently strongly to suggest that population changes represent a driving force behind the evolution of education and human capital at the Prussian-Russian border. Compared to the baselines reported in column 1 of Table 1.10, the point estimate of the partition effect changes by less than 10 pp when adding the various population shares (columns 2 and 3). Further, the tendency of the partition effect to decline over time, with the small rebound in 1931/32, is hardly affected.

The measure of population replacement in 1921 has no effect in 1911, validating that it indeed captures more recent population movements. The estimates presented in column 3 suggest that these movements have a strong negative effect on primary enrollment in the years 1920/21 and 1925/26. Keeping in mind that German emigration encompassed the emigration of a considerable number of German teaching personnel, these results seem reasonable: In counties where population replacement took place to a relative intensive degree, the Polish takeover of the educational infrastructure after WWI encountered more frictions on average in the short run than elsewhere because Polish teaching personnel had to be recruited first. However, these frictions appear to have been overcome by the year 1931/32, as the population replacement effect is insignificant then. In fact, controlling for all three population measures suggests that the Prussian partition effect is not statistically different from zero anymore in 1931/32. The population composition controls affect neither size nor significance of the estimated partition effects at the Austrian-Russian border (see Table 1.11), thereby confirming the presumption that at least before WWII, the population has been very persistent along this border.

Table 1.10: Enrollment at Prussian-Russian border 1911-1931, population controls

	(1)	(2)	(3)
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Prussian Side = 1	0.832***	0.847***	0.839***
	(0.015)	(0.027)	(0.027)
Share Jewish		0.368	0.374
		(0.260)	(0.266)
Share Protestant		0.042*	0.022
		(0.022)	(0.042)
Share Born Outside Partition 1921			0.083
			(0.114)
Observations	54	54	54
R-squared	0.997	0.997	0.997
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Prussian Side = 1	0.378***	0.346***	0.454***
	(0.044)	(0.076)	(0.097)
Share Jewish		-0.458	-0.429
		(0.784)	(0.766)
Share Protestant		0.076	0.307*
		(0.157)	(0.175)
Share Born Outside Partition 1921			-0.851***
			(0.309)
Observations	54	54	54
R-squared	0.902	0.904	0.916
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Prussian Side = 1	0.066*	0.034	0.163***
	(0.033)	(0.049)	(0.058)
Share Jewish		-0.582	-0.492
		(0.520)	(0.485)
Share Protestant		0.029	0.346*
		(0.257)	(0.197)
Share Born Outside Partition 1921			-0.945***
			(0.220)
Observations	54	54	54
R-squared	0.647	0.657	0.752
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Prussian Side = 1	0.098***	0.092***	0.069
	(0.026)	(0.030)	(0.045)
Share Jewish		-0.108	-0.134
		(0.407)	(0.429)
Share Protestant	0.029	-0.032	
		(0.183)	(0.218)
Share Born Outside Partition 1921			0.158
			(0.191)
Observations	54	54	54
R-squared	0.550	0.550	0.555
Geo, City Controls	Yes	Yes	Yes
2nd Order Polynomial	Yes	Yes	Yes

Notes: Two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.11: Enrollment at Austrian-Russian border 1911-1931, population controls

	(1)	(2)	(3)
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Austrian Side = 1	0.412***	0.427***	0.423***
	(0.043)	(0.049)	(0.049)
Share Jewish		0.391	0.422
		(0.374)	(0.299)
Share Protestant		3.903	3.748
		(2.351)	(2.404)
Share Born Outside Partition 1921			-1.9
			(1.289)
Observations	44	44	44
R-squared	0.930	0.939	0.944
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Austrian Side = 1	0.156***	0.123***	0.117**
	(0.045)	(0.044)	(0.043)
Share Jewish		-1.058***	-1.291***
		(0.315)	(0.298)
Share Protestant		-0.550	0.148
		(4.774)	(4.481)
Share Born Outside Partition 1921			1.805*
			(0.967)
Observations	44	44	44
R-squared	0.870	0.898	0.907
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Austrian Side = 1	0.028	0.031	0.033
	(0.050)	(0.054)	(0.055)
Share Jewish		-0.177	0.036
		(0.688)	(0.707)
Share Protestant		-2.593	-2.720
		(4.976)	(5.215)
Share Born Outside Partition 1921			-1.390
			(1.793)
Observations	44	44	44
R-squared	0.428	0.434	0.445
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Austrian Side = 1	-0.009	-0.004	-0.004
	(0.027)	(0.028)	(0.028)
Share Jewish		0.094	0.061
		(0.264)	(0.277)
Share Protestant		-1.525	-1.592
		(3.007)	(3.028)
Share Born Outside Partition 1921			0.230
			(0.459)
Observations	44	44	44
R-squared	0.662	0.668	0.670
Geo, City Controls	Yes	Yes	Yes
2nd Order Polynomial	Yes	Yes	Yes

Notes: Two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

With regard to literacy, the importance of the population composition appears to be even lower. The partition effect only rises somewhat considerably at the Prussian-Russian border in 1931 when controlling for the share of the population born outside of the respective residential partition (column 3 of Table 1.12), indicating a depressing effect of the recent population replacement on human capital, possibly via the former's (short-run) negative effect on enrollment previously detected. Again, the partition effect at the Austrian-Russian border is predominantly unaffected (Table 1.13).

Table 1.12: Literacy at Prussian-Russian Border 1921-1931, population controls

	(1)	(2)	(3)
Dep. Variable	Share of Literates 1921	Share of Literates 1921	Share of Literates 1921
Prussian Side = 1	0.244***	0.263***	0.280***
	(0.012)	(0.019)	(0.024)
Share Jewish		0.401	0.406
		(0.251)	(0.252)
Share Protestant		0.070	0.107
		(0.059)	(0.066)
Share Born Outside Partition 1921			-0.135
			(0.131)
Observations	54	54	54
R-squared	0.967	0.969	0.970
Mean on Russian Side	0.509	0.509	0.509
Dep. Variable	Share of Literates 1931	Share of Literates 1931	Share of Literates 1931
Prussian Side = 1	0.192***	0.213***	0.245***
	(0.011)	(0.013)	(0.017)
Share Jewish		0.433***	0.468***
		(0.157)	(0.165)
Share Protestant		0.035	0.119*
		(0.107)	(0.070)
Share Born Outside Partition 1921			-0.218***
			(0.067)
Observations	54	54	54
R-squared	0.976	0.979	0.982
Mean on Russian Side	0.517	0.517	0.517
Geo, City Controls	Yes	Yes	Yes
2nd Order Polynomial	Yes	Yes	Yes

Notes: Two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.13: Literacy at Austrian-Russian border 1921-1931, population controls

	(1)	(2)	(3)
Dep. Variable	Share of Literates 1921	Share of Literates 1921	Share of Literates 1921
Austrian Side = 1	0.170***	0.170***	0.171***
	(0.042)	(0.046)	(0.048)
Share Jewish		0.298	0.332
		(0.259)	(0.245)
Share Protestant		2.403*	2.301*
		(1.291)	(1.303)
Share Born Outside Partition 1921			-0.264
			(0.787)
Observations	44	44	44
R-squared	0.908	0.915	0.915
Mean on Russian Side	0.462	0.462	0.462
Dep. Variable	Share of Literates 1931	Share of Literates 1931	Share of Literates 1931
Austrian Side = 1	0.093***	0.103***	0.103***
	(0.027)	(0.030)	(0.032)
Share Jewish		0.366	0.427*
		(0.228)	(0.250)
Share Protestant		-1.244	-1.119
		(1.749)	(1.561)
Share Born Outside Partition 1921			-0.430
			(0.506)
Observations	44	44	44
R-squared	0.899	0.908	0.911
Mean on Russian Side	0.491	0.491	0.491
Geo, City Controls	Yes	Yes	Yes
2nd Order Polynomial	Yes	Yes	Yes

Notes: Two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

1.6.2 Supply and endowment of schools

Given the observed fading discontinuities in primary school enrollment at both borders, it is natural to ask how far this process was accompanied or had even been preceded by an alignment of the network of schools across the former partitions. Therefore, I calculate the number of primary schools per 1000 inhabitants of primary school age for each county from 1911 to 1960/61 in order to test for discontinuities in this proxy for school supply and density.[5] In order to adjust for the fact that the schools in the Russian partition were very ill-equipped compared to the ones in the Austrian and Prussian partitions respectively, I also calculate the number of classes in primary schools per 1000 inhabitants of primary school age. 'Classes' proxy for school quality in this context because they denote how many different grades of a school are taught within the same classroom. Finally, I also compute the number of primary school teachers per 1000 school-aged inhabitants as an additional measure of school endowment. A drawback is that data on the number of classes and teachers are not available for the school years 1931/32 and 1960/61.

Unsurprisingly, the results displayed in Panel A of Table 1.14 suggest that the network of

[5] I use the number of inhabitants of primary school age because the number of primary school students is endogenous to the supply of primary schools.

primary schools in the Prussian partition was both denser and of better quality compared to the Russian partition in 1911/12: The average number of primary schools per 1000 school-age children jumps by more than five at the Prussian side of the border or by about 250%. The discontinuities in the provision of classes and teachers is of even greater magnitude; it amounts to about 700% and 450% respectively, underscoring the woeful state of primary education in the Russian partition at that time.

During the early years of the Second Republic, the estimates of the partition effect decrease considerably in magnitude, as the average numbers of primary schools and classes triple in the former Russian partition, with the number of teachers growing at an even higher rate. However, the estimates neither turn statistically insignificant nor quantitatively negligible, with only minor changes in 1925/26. Still, the supply of teachers and classes increases steadily in the former Russian partition, indicating a convergence of these two quality measures to a relatively high level of standard. The difference with regard to the density of schools remains sizable in 1931/32, as it actually shows a slight rebound compared to the previous post-WWI periods. Only in 1960/61, with the former Russian side of the Prussian-Russian partition border providing on average seven primary schools per 1000 children at school-age, the point estimate of the former Prussian partition turns negative and weakly significant.

Interestingly, there is no discontinuity in school density at the Austrian-Russian border in 1911/12. The results presented in Panel B of Table 1.14 are rather inconclusive with regard to this measure across time, indicating a small advantage on the territory of the former Austrian partition only in 1931/32. Nevertheless, in 1960/61, the effect is small, negative and weakly statistically significant, similar to the Prussian partition effect in the same time period.

Still, the Austrian side of the border exhibits a positive difference in school quality as measured by the number of classes and teachers in 1911/12. Thus, the advantage of the Austrian partition over the Russian partition appears to run through the quality and not the number of schools close to the end of the imperial rule. This advantage, however, is much smaller than that of the Prussian partition. It declines over time, with the effect on teacher supply being insignificant in 1931/32, while the estimated coefficient is still positive and sizable. But whether or not the disappearance of statistical significance is due to imprecise estimates, the supply of both classes and teachers increases on either side of the Austrian-Russian border to a similar level as along the Prussian-Russian counterpart.

Table 1.14: Schools, classes, teachers at partition borders 1911-1961

	(1)	(2)	(3)
Panel A: Prussian-Russian Border			
Dep. Variable	Schools Per 1000 School-Aged 1911	Classes Per 1000 School-Aged 1911	Teachers Per 1000 School-Aged 1911
Prussian Side = 1	5.408***	16.559***	11.305***
	(0.535)	(0.591)	(0.642)
Observations	54	54	54
R-squared	0.936	0.981	0.977
Mean on Russian Side	2.143	2.262	2.413
Dep. Variable	Schools Per 1000 School-Aged 1921	Classes Per 1000 School-Aged 1921	Teachers Per 1000 School-Aged 1921
Prussian Side = 1	1.684**	7.894***	3.463***
	(0.632)	(0.891)	(0.741)
Observations	54	54	54
R-squared	0.740	0.810	0.603
Mean on Russian Side	6.659	7.915	9.377
Dep. Variable	Schools Per 1000 School-Aged 1926	Classes Per 1000 School-Aged 1926	Teachers Per 1000 School-Aged 1926
Prussian Side = 1	1.820***	7.332***	4.781***
	(0.450)	(1.164)	(0.976)
Observations	54	54	54
R-squared	0.876	0.714	0.590
Mean on Russian Side	6.645	13.727	13.177
Dep. Variable	Schools Per 1000 School-Aged 1932		
Prussian Side = 1	2.374***		
	(0.575)		
Observations	54		
R-squared	0.818		
Mean on Russian Side	6.042		
Dep. Variable	Schools Per 1000 School-Aged 1961		
Prussian Side = 1	-1.595*		
	(0.917)		
Observations	53		
R-squared	0.868		
Mean on Russian Side	7.024		
Panel B: Austrian-Russian Border			
Dep. Variable	Schools Per 1000 School-Aged 1911	Classes Per 1000 School-Aged 1911	Teachers Per 1000 School-Aged 1911
Austrian Side = 1	-0.089	4.998***	4.500***
	(0.574)	(0.959)	(0.948)
Observations	44	44	44
R-squared	0.448	0.879	0.884
Mean on Russian Side	2.47	2.648	3.144
Dep. Variable	Schools Per 1000 School-Aged 1921	Classes Per 1000 School-Aged 1921	Teachers Per 1000 School-Aged 1921
Austrian Side = 1	-2.338**	0.852	2.196**
	(0.865)	(0.863)	(0.994)
Observations	44	44	44
R-squared	0.580	0.609	0.838
Mean on Russian Side	5.71	8.154	9.01
Dep. Variable	Schools Per 1000 School-Aged 1926	Classes Per 1000 School-Aged 1926	Teachers Per 1000 School-Aged 1926
Austrian Side = 1	-0.550	3.152**	2.020
	(0.563)	(1.499)	(1.464)
Observations	44	44	44
R-squared	0.560	0.675	0.644
Mean on Russian Side	5.234	12.435	12.392
Dep. Variable	Schools Per 1000 School-Aged 1932		
Austrian Side = 1	1.06**		
	(0.471)		
Observations	44		
R-squared	0.692		
Mean on Russian Side	4.751		
Dep. Variable	Schools Per 1000 School-Aged 1961		
Austrian Side = 1	-0.881*		
	(0.500)		
Observations	59		
R-squared	0.838		
Mean on Russian Side	6.024		
City, Geo Controls	Yes	Yes	Yes
2nd Order Polynomial	Yes	Yes	Yes

Notes: Two-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

1.7 Conclusion

This paper examines the importance of history in the aftermath of a large-scale natural experiment that had split the Polish nation into three partitions under different imperial rule. Beginning in the last years of existence of the three empires, I find that the educational differences among the partition populations that the empires had instilled during their century-long control over Poland were tremendous: The Prussian partition surpassed the Russian one in terms of primary enrollment by more than 80 pp, reflecting both the high standard and enforcement of public education in Prussia. The differences in enrollment between the Austrian and the Russian partition, in turn, were less dramatic, but still amounted to more than 40 pp. However, my results further show that these gaps have essentially been eliminated by 1931 or less than 15 years since the reconstitution of the Polish state after WWI. They do not reappear after the chaos of WWII, but instead remain nonexistent during the communist period of Polish history. The achievement of universal literacy took longer to reach due to simple demographic reasons, but it was essentially completed as well in 1960. It is reasonable to conclude therefore that the non-persistence of imperial legacies in education observed by Grosfeld & Zhuravskaya (2015) in present-day Poland was accomplished relatively soon after the Poles had been fully handed back the control over their educational system.

I further provide evidence that population changes in the partitions, in particular the emigration of the German minority from the former Prussian partition in interwar Poland, had at best short-run negative effects on education. The equalization of enrollment, in turn, was accompanied by an expansion of both the network of primary schools and the quality of schools as proxied by classes and teachers. Thus, the public supply of educational facilities was successfully directed at the previously disadvantaged former Russian partition in central Poland.

There are still numerous open questions in the context of the history of the partitions and education in Poland. I show that within a relatively short episode of this history, the imperial legacies on the extensive margin of primary education fade out. However, it could well be the case that legacies continue to exist not only on the intensive margin of quality (Bukowski, 2016), but also on higher levels of educational attainment. Further, I cannot provide direct evidence on the importance of educational institutions and culture in the time frame of my sample due to the lack of disaggregated institutional data and, even more, the lack of survey information that allows eliciting the role of social norms and culture as utilized by Grosfeld & Zhuravskaya (2015) and Bukowski (2016) in contemporary Poland. However, if there were imperial legacies in educational institutions and culture, then they did not restrain (or advance) one partition over another permanently, at least in terms of basic human capital. Another interesting question concerns the importance of spillovers after the dissolution of the imperial borders. For example, it is conceivable that the counties in the former Russian partition that were located close to the former Austrian border were able to adopt elements of the already advanced system of school organization from the former Austrian partition after 1918. Answering this question is currently beyond the depth of my data.

In total, the case of human capital formation in the aftermath of the Partitions of Poland suggests that substantial differences in education that were laid down over the course of a century can be leveled within a modern state in a comparatively short period of time. The finding that Poland did indeed accomplish this is, of course, to be interpreted within the historical context of my sample period: After its regained independence after WWI, Poland was competing with European nations, in particular Germany, that were already much more advanced in terms of universal education. The situation might have been somewhat comparable to that of Prussia and its competition with England and France one hundred years earlier that stimulated large-scale and swift reforms of the Prussian education system. Noteworthy, my finding of rapid integration in educational terms corresponds to the assessment of Trenkler & Wolf (2005) that by the middle of the 1920s, Poland's internal markets were already well-integrated, too.

Finally, the richness of the school data that is not nearly fully exploited by this study calls for further research with regard to the potential (non-)disappearance of other educational legacies such as along the gender or religious dimension.

Chapter 2

Ethnic Favoritism Revisited: Competitive Voting in Ghana

2.1 Introduction

Ethnic favoritism has become an intensely studied phenomenon in the economic analysis of both governance in general and democracy in particular. Needless to say, investigations into ethnic favoritism have been largely directed at countries and continents where ethnic identities are considered to be strong and relevant. For example, Burgess et al. (2015) show that ethnic favoritism, as measured in terms of road investments in Kenyan districts that are co-ethnic with the respective political ruler, has been widespread under various autocratic rulers of different ethnic affiliation, while it appears to have been curbed after Kenya's transition to democracy in the early 1990s. Kramon & Posner (2016), in turn, assess that with regard to education, ethnic favoritism in Kenya has all but disappeared under democracy. In the broader context of sub-Saharan Africa, Franck & Rainer (2012) estimate large and widespread ethnic gains in terms of education and health originating from time periods when an ethnic group has been co-ethnic with its respective country's leader. On average, these gains remain unaffected by whether the form of government is democratic or autocratic. De Luca et al. (2016) document widespread favoritism in terms of nighttime lights on the level of ethnic homelands. Using a global dataset, they suggest that ethnic favoritism is common across, but not limited to sub-Saharan Africa, with rather negligible dampening effects of the quality of political institutions.

Under which circumstances could ethnic favoritism actually occur in a democratic system of governance? As pointed out by Amodio & Chiovelli (2016), democracy can broaden the scope for strategic interactions between politicians and ethnic leaders. The mixed evidence on the prevalence of ethnic favoritism under democracy suggests that the outcomes of these interactions might be heterogeneous with regard to the extent of ethnic favoritism that they are able to provide (or prevent). From the perspective of economic theories of democracy, the empirical prevalence of ethnic favoritism under democracy can be related to the ‘core’ voter concept of Cox & McCubbins (1986): Political parties differ with regard to their ability to redistribute towards different groups among the electorate, while the groups in turn differ strongly in terms of their ideological party preference. In equilibrium, this results in each party focussing its redistributive efforts on the specific group(s) to which it can redistribute the easiest. Hence, groups are generally not courted by more than one party, making them solid, delimited political blocks which enjoy high patronage as soon as ‘their’ party climbs to power.

At first sight, the case of the West African country of Ghana bears a large similarity to other nations whose democratization has been the subject of previous studies: Ghana returned to constitutional democracy in 1992 in the course of the so-called third wave of democratization (Huntington, 1991) after decades of alternating democratic turmoil and military rule. The latter culminated in the eleven-year reign of Jerry John Rawlings as the Chairman of the Provisional National Defence Council. Similar to the Kenyan experience, Rawlings then formed a political party, the New Democratic Congress (NDC), and went on to become the first democratically elected president of the Ghanaian fourth republic. All quadrennial elections since 1992 have been considered free and fair; they have also resulted in the first peaceful transition of power from the NDC to its main challenger, the New Patriotic Party (NPP), in the year 2000 when the constitution barred Rawlings from running for a third presidential term. From the viewpoint of ethnicity, Ghana’s ethnic landscape is as much a product of the arbitrariness of colonial borders as many other African nations. The Asante, who have once been the rulers of the powerful Ashanti empire, exhibit a strong and long-lived ethnic affiliation to the NPP. The NDC, in turn, has strong ties to the Ewe ethnic group, given that Jerry John Rawlings belongs to this group as well.

However, it is questionable whether Ghana’s ethnic setup features the necessary preconditions for large-scale redistribution towards the respective president’s co-ethnics because none of the two politically invested groups is actually large enough to secure a majority of the national vote by its own. This suggests that the two political parties rather have to compete for the votes of the more unaffiliated Ghanaian electorate outside of the parties’ ethnic boundaries. This political constellation bears more similarity to the probabilistic or ‘swing’ voting models pioneered by Lindbeck & Weibull (1987) and Dixit & Londregan (1996). The essential prediction of these models is that those groups that contain the highest share of non-partisan, ‘moderate’ voters will be promised the highest share of redistributive transfers by the political parties because the moderates are the most likely to ‘swing’ their votes from one party to the other in return for economic remuneration. Correspondingly, survey evidence in Lindberg & Morrison (2005) and Lindberg (2012) indicates a relatively high and growing share of swing voters among the Ghanaian electorate. Evidence by Banful (2011) on the political economy of local budget allocations by the Ghanaian central government further suggests that districts which exhibit tighter vote margins in a presidential election receive higher allocations afterwards. Banful (2011) does not explore any ethnic dimensions of these voting patterns though.

This paper attempts to broaden the scope for the understanding of ethnic favoritism. It does so by exploiting electoral results and changes of government in Ghana between 1992 and 2008. I first show that the two Ghanaian ethnicities that are co-ethnic with the varying presidents become economically *worse off* in relative terms soon after the country's return to democracy. I then test the prediction of the probabilistic voting theory that close voting should be associated with economic transfers if there are moderate groups of voters to be swayed. I find that there is indeed a positive association between close voting and economic prosperity, thereby confirming the finding of Banful (2011). However, I further show that this association runs entirely through the homeland of the large, politically unaffiliated ethnic group of the Akan, while it is not detectable with regard to other ethnicities. Taken together, these results suggest that while the eagerness of political parties to form multi-ethnic electoral coalitions has the effect of constraining ethnic favoritism towards the co-ethnics of the respective political leader, the same eagerness can give rise to ethnic favoritism directed towards groups that signal their readiness to be courted by the political contestants.

The paper hence contributes to the aforementioned literature on ethnic favoritism in democracies. It further makes a methodological contribution to the literature: Because there are no official statistics on economic prosperity at the district level, which is the level of observation in the following, I make use of nighttime lights as a proxy variable. While nighttime lights have already been extensively utilized in the context of ethnic favoritism, I reaffirm the usefulness of nighttime lights for detecting patterns of favoritism at the sub-national level.

I present an abridgment of the probabilistic voting theory, as well as more background information on Ghana's system of governance, ethnic divisions, and elections in Section 2.2. Section 2.3 presents the data for the intended analysis, followed by the empirical strategy in Section 2.4. Section 2.5 first provides the results on the negative effects of co-ethnicity on economic prosperity and then moves on to point out the positive relationship between the economic prosperity of the Akan-dominated districts and their close-voting behavior. The subsequent Section 2.6 provides a discussion of the results. Section 2.7 concludes with a prospect for further research on democracy in Ghana.

2.2 Background

2.2.1 Probabilistic voting theory

The two seminal theoretical contributions to the understanding of intense political competition for similar groups of voters are Lindbeck & Weibull (1987) and Dixit & Londregan (1996). The formulation of the latter is more general in the sense that it also features an outcome in which each party targets its loyal electoral core instead of the unaffiliated voters are the prime recipients of redistribution. However, as this outcome is not particularly relevant for the Ghanaian case, I omit it.

The models' setups are very similar to each other: Following the notation of Lindbeck & Weibull (1987), two political parties, A and B, compete for the electorate I consisting of n voters indexed by i. Voters receive an exogenous and fixed gross income of $\omega = (\omega_1, \ldots, \omega_n) > 0$. The electorate is further divided into m disjoint and distinguishable groups, where the kth group I_k contains n_k voters (with $\sum n_k = n$). Each member of group k receives a transfer z_k, whereas the sum of transfers across all groups must satisfy the balanced-budget constraint $\sum n_k z_k = 0$. $x_{k(i)}$ then denotes a feasible balanced-budget redistribution that individual i receives as a member of group k. Voters are rational and derive utility u_i from two (additively-separable) components: $v_i(\omega_i + x_{k(i)})$, which exhibits strictly positive, but strictly decreasing marginal returns to consumption, and their ideological preferences a_i and b_i for other policies enacted by party A respectively B. Party A can convince a voter with a stronger ideological preference for party B to vote nevertheless for party A by offering her such a generous transfer that her utility from consuming the transfer exceeds the utility from voting in accordance with her political beliefs for party B. However, parties cannot observe individual ideological preferences, which is why they assign probability distributions F_i over them (hence the term 'probabilistic voting') with the corresponding density functions f_i. Under the assumption that f_i is unimodal and symmetric, a single scalar α_i then denotes the expected party bias of an individual for party A over party B. In equilibrium, both parties will follow symmetric strategies and favor those groups in the electorate with a relatively high share of moderate individuals with weak partisan biases. Supposing further that all individuals have the same consumption preferences and party preference distributions differ only between, but not within groups, the per-capita transfer to a group is then a decreasing function of the absolute value of the expected party bias in the group:

$$\sum v'(c_i)/n_k = \lambda/f(\alpha_i) \tag{2.1}$$

where $i \in I_k$ and λ is the Lagrangian multiplier from the vote maximization problem of the political parties. The density function on the right-hand side is decreasing in the absolute expected party bias, while the marginal utility of consumption on the left-hand side is decreasing in the level of consumption. Hence, in equilibrium, lower values of the expected party bias translate into higher transfers and the votes of the groups with the lowest expected partisan biases are the most contested, resulting in them swinging between the two parties.

Lindbeck & Weibull (1987) and Dixit & Londregan (1996) further point to a number of special cases: If the party preference distributions are the same for all individuals, then the political equilibrium is identical to the outcome of maximizing the utilitarian social welfare function subject to the balanced-budget constraint. Moreover, if there are no party biases in the population, then the expected vote share for each party is exactly 50%. The assumption of rational voters further implies that all promises of balanced-budget redistribution are credible and that voters will not be systematically fooled by political parties promising them positive redistribution, but then not enacting it accordingly.

A straight-forward prediction of the probabilistic voter models is that groups which comprise a high share of politically unaffiliated swing voters should on average receive positive transfers in the aftermath of an election. Because I observe neither voters nor their political affiliations directly, I assume that the observed ex post electoral competitiveness of a district, calculated as the absolute vote margin between the two relevant parties, reflects the underlying distribution of partisan bias. This translates into estimating a reduced-form relationship between a measure of economic transfers and the vote margin at the district level.

2.2.2 Democratic institutions and elections in Ghana

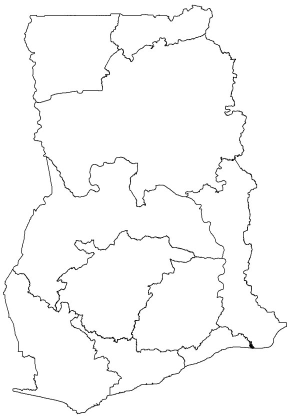
After eleven years of authoritarian rule under the leadership of Flight Lieutenant Jerry John Rawlings, which had already been preceded by a series of military *coups d'état*, Ghana returned to democracy as the country's system of government with the adoption of a new constitution by referendum in April 1992. Since then, the candidate who wins the majority of the national vote in a presidential election becomes the next president. Rawlings emerged as the winner of the first presidential election held later in 1992 on the ticket of the newly-formed National Democratic Congress (NDC). In accordance with the constitutional provisions, elections have since then been held every four years, with Rawlings winning a second term in 1996. However, because the constitution limits the number of presidential terms to two, Rawlings could not run again in 2000. His designated successor John Atta Mills was defeated by the candidate of the then-oppositional New Patriotic Party (NPP), John Kufuor, marking the first democratic transfer of power under the 1992 constitution. Kufuor went on to win a second term in 2004, while his designated successor Nana Akufo-Addo in turn lost to Atta Mills in the very tightly contested 2008 election, thereby returning power to the NDC. However, Atta Mills unexpectedly died in office shortly before the completion of his first term. After the smooth transfer of power to the by-then vice president John Mahama, the latter went on to win the 2012 election, only to be defeated by Akufo-Addo and the NPP in the latest election in 2016.

Since 1996, parliamentary elections have been held on the same date as the presidential ones. While split-ticket voting with regard to presidential and parliamentary candidates, or 'skirt-and-blouse' voting, as it is commonly called in Ghana, is possible, so far every victorious presidential candidate's party has also won the majority of the parliamentary seats. In addition, the parliament's position is generally considered weak in comparison to the office of the president. For example, allocations of public funds are largely controlled by the president respectively his appointees (Banful, 2011), which is why I follow the literature and focus exclusively on the presidential elections in the following.

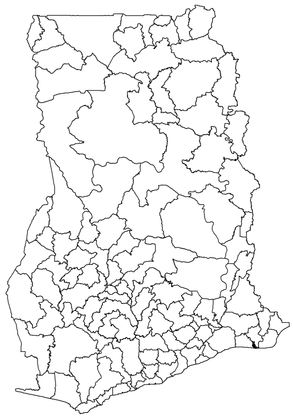
Between 1992 and 2000, voting took place in 200 electoral constituencies, while this number was increased to 230 in 2004 and further to 275 in 2012. Every constituency is comprised within the boundaries of a district assembly. One district assembly may comprise several constituencies, but not vice versa. District assemblies (hereafter referred to as districts) constitute the lowest level of administrative divisions in Ghana, with ten regions being located between the national and the district level. While the regions are relatively unimportant as institutions of governance, the districts are recipients of relevant funds from the central government (Banful, 2011). In parallel with the state of constituencies, the district count was kept constant at 110 between 1992 and 2000. By 2004, it has been raised to 138 by splitting a number of districts in two, allowing it to restore the original 110 boundaries. However, between 2004 and 2008, the district count was further increased to 170 not only by splitting, but also by redrawing a number of districts, making it impossible to aggregate them back to their former shape. Another 46 districts have been added since 2012, raising the current count to 216. Due to these administrative changes, I focus the analysis on the elections 1992-2004 in the following. Outlines of the ten regions and 110 districts provided by IPUMS International (2017) are displayed in Figure 2.1.

In general, with the exception of 1992, the two parties NDC and NPP together have always acquired more than 90% of the national vote. If neither party wins more than 50% of the national vote in the first round of an election, which is possible due to a handful of small parties competing with the two larger ones in the first round, a second or runoff round is held. Within the sample period, this happened only in 2000, with the NPP expanding its lead on the NDC from the first to the second round.

Figure 2.1: Administrative divisions of Ghana



(a) 10 Regions



(b) 110 Districts

2.2.3 Ethnic politics in Ghana

Both NDC and NPP exhibit strong ethnic ties to different ethnic groups in Ghana. While Rawlings' father was British, his mother belonged to the Ewe people mainly located in Ghana's Volta region along the border to Togo. Due to Rawlings' role as the founding father and long-term chairperson of the NDC, the party usually obtains very high vote shares in the Ewe-dominated territories. Conversely, the NPP is strongly associated with the Asante people, whose homeland is the Ashanti region. This association even predates the 1990s, with the Asante already supporting predecessors of the NPP which stood in opposition to Ghana's first democratically elected, but increasingly authoritarian president Kwame Nkrumah in the 1960s (Ichino & Nathan, 2013). In terms of their origin and language, the Asante are a subgroup of the Akan people who form the largest (45% including the Asante) ethnic group within Ghana.[1] According to national population accounts, the Asante themselves represent about 15% of the population, while the Ewe represent about 12%. Thus, the two groups are almost equally strong in terms of members, but each of them is far from being able to achieve a majority of the national vote on its own.

The Akan homelands are predominantly located in the regions surrounding Ashanti, namely Brong-Ahafo to the north, the Western region to the south-west, and the Eastern region to the south-east. The Central region to the south of Ashanti comprises the homeland of the Fante, the largest identifiable Akan group after the Asante. The remaining ethnic groups of Ghana, who comprise about 45% of the population, are neither closely related to the Asante, nor to the Akan in general, nor to the Ewe. The coastal homeland of the Ga-Adangbe people (7% of the population) contains Ghana's capital Accra and is therefore densely populated, as well as increasing ethnically mixed due to the steady

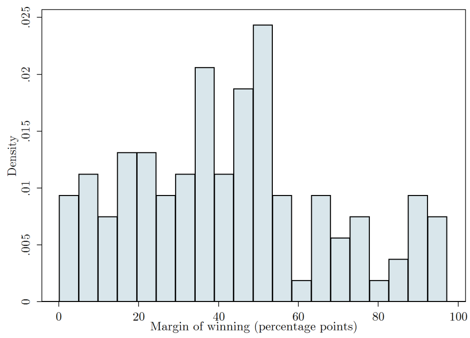
[1] 'Akan' in the following refers to the Akan groups other than the Asante.

influx of internal migrants from other parts of the country. The ethnic groups that inhabit the north of Ghana, most notably the Mole-Dagbani (15% of the population) and various spatially more concentrated groups like the Guan (4%), the Gurma (3%) and the Grusi (2%) are particularly distinct from the other ethnicities with regard to their indigenous language and their religion, which is predominantly Islam. Districts where the population share of the respective ethnic group exceeds 50% are displayed in Figure B.1.

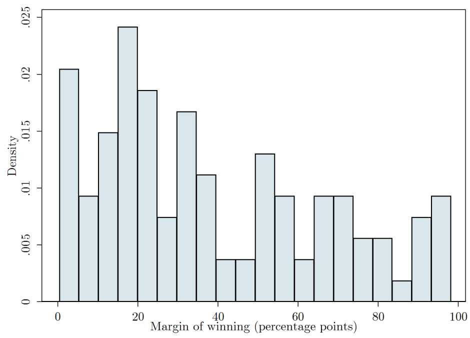
2.2.4 Competitiveness of Ghanaian elections

There are strong indications that elections in Ghana have become much more competitive in terms of vote margins since the return to democracy in 1992. This tendency can be observed both at the national and at the district level: The national vote margin steadily declines from 28.1 pp in 1992 over 17.7 pp in 1996 to 13.8 pp in 2000 and further to 7.81 pp in 2004. The district-level margins of winning, as displayed in the histograms of Figure 2.2, visibly tightens in particular from 1992 to 2000. Complementary, in Figure 2.3, the district-level vote margins of two consecutive elections are plotted against each other. Observations below the 45° line indicate districts which become more competitive from one election to the next one. While some districts also become less competitive over time, plotting the vote margins of the 2004 against those of the 1992 elections illustrates that the majority of the district-level observations moves towards tighter electoral outcomes.

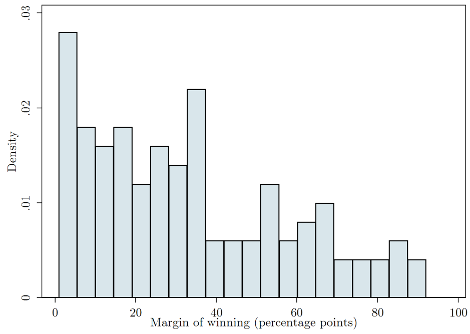
Figure 2.2: Histograms of vote margins in presidential elections 1992-2004



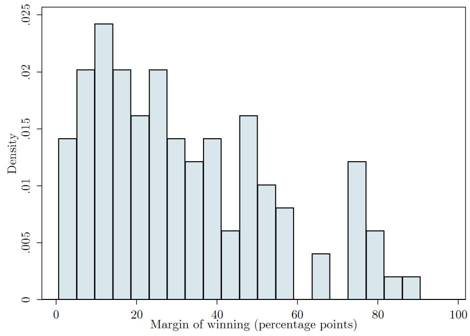
(a) 1992



(b) 1996

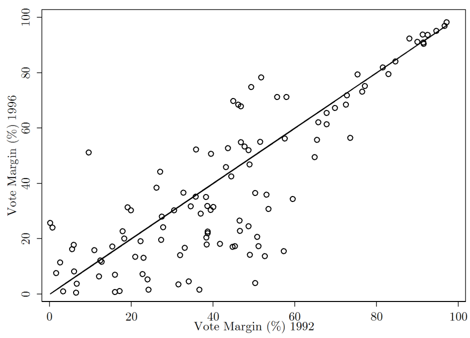


(c) 2000

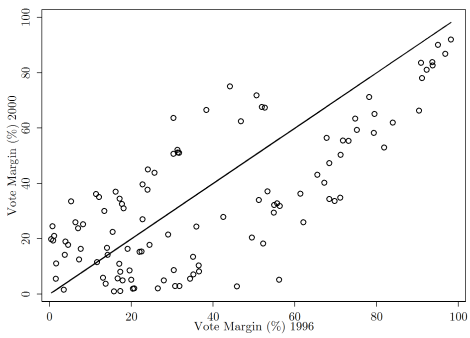


(d) 2004

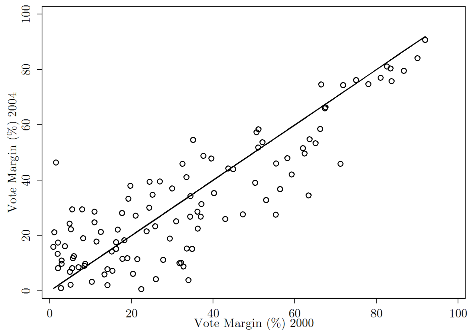
Figure 2.3: Margins of victory in presidential elections 1992-2004



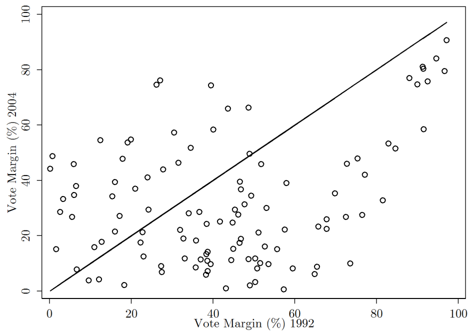
(a) 1992 vs. 1996



(b) 1996 vs. 2000



(c) 2000 vs. 2004

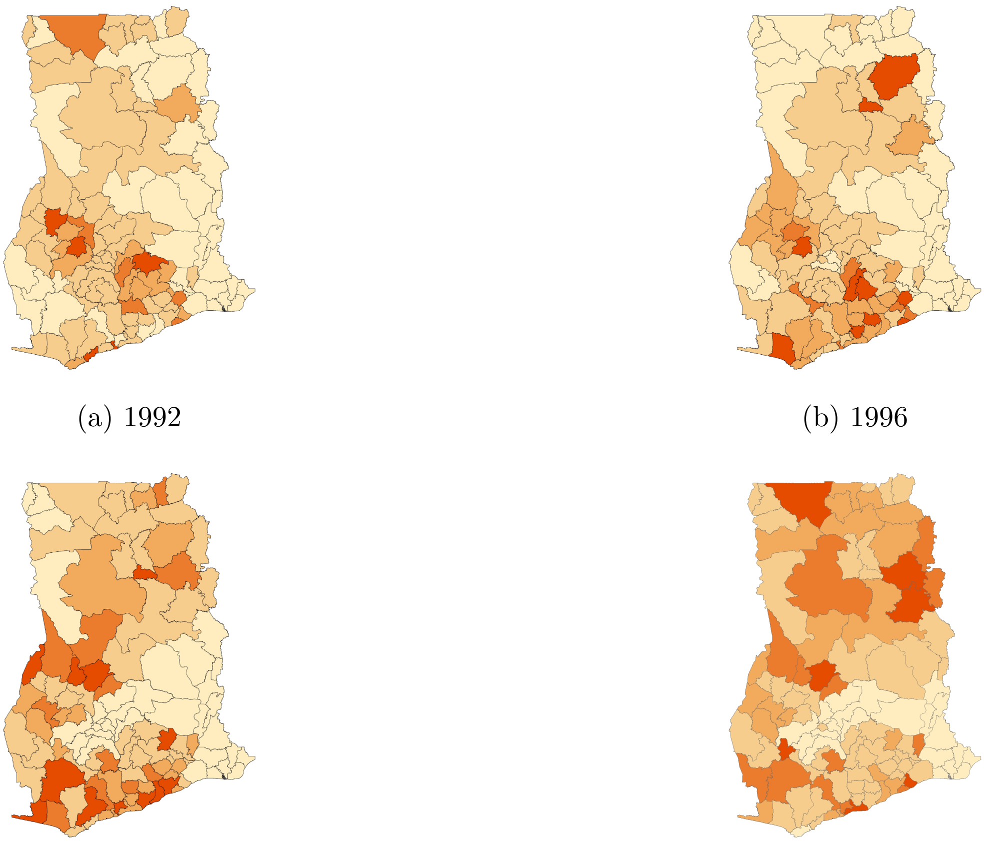


(d) 1992 vs. 2004

These descriptives are consistent with representative survey evidence by Lindberg & Morrison (2005): In 2003, 18% of the Ghanaian respondents fall into the category of swing voters. This classification is based on whether respondents stated that they had voted consistently for one party in the 2000 and 1996 elections or whether they had switched between parties.

Figure 2.4 displays the evolution of the spatial distribution of competitive districts across the elections from 1992 to 2004. The five different degrees of coloring reflect five different bins of vote margins: up to 5, 5-10, 10-20, 20-50, and 50- percentage points, with the color intensity increasing with the shrinking vote margin. The maps suggest that highly competitive districts are a rare phenomenon in the early years of the reconstituted Ghanaian democracy. By the time of the 2000 elections, however, they are well-established in the Akan-dominated Central, Western, and Brong-Ahafo regions surrounding the Ashanti region, as well as in the small, but densely populated Greater Accra region. The Ashanti and Volta regions, i.e. the ethnic homelands of the Asante respectively Ewe people, are unsurprisingly not closely contested in any election.

Figure 2.4: Spatial distribution of vote margins in presidential elections 1992-2004



(a) 1992

(b) 1996

(c) 2000

(d) 2004

2.3 Data

Information on the spatial distribution of ethnic groups at the district level is obtained from the 10% extract of the Ghanaian Population and Housing Census of the year 2000 provided by IPUMS International (2017). While the census lists a total of 55 ethnic identities, it also subsumes the latter under eight aggregate labels. Given that more than half of the more disaggregated group classifications each represent less than 1% of the Ghanaian population, the aggregate labels are used in the following.[2]

Electoral data on the constituency level for most elections are provided by the Electoral Commission of Ghana, while some results are supplemented from online news repositories. Because the boundaries of constituencies are always fully comprised within the district boundaries, the votes of multiple constituencies within the same district can be summed up in order to compute the respective vote shares at the district level. The 200 electoral constituencies of the period 1992-2000 and the 230 constituencies in 2004 are matched into the 110 original districts, resulting in a balanced panel dataset spanning all elections from 1992 to 2004. While this procedure ignores the aforementioned split-up of some districts between the years 2000 and 2004, Banful (2011) provides evidence that the probability whether a district was split up is unrelated to its past electoral performance.

Due to the absence of repeated district-level data on poverty, GDP and the like, a district's economic prosperity is proxied with its intensity of nighttime lights. The stable lights composite of the Version 4 DMSP-OLS Nighttime Lights Time Series provided by the National Oceanic and Atmospheric Administration (NOAA)[3] allows the yearly computation of the logarithm of the nighttime lights per pixel at the district level from 1993 to 2008. This timing is due to the fact that Ghanaian elections always take place in the month of December of an election year, resulting in the new government assuming office not before the beginning of the following year. A presidential term is thus considered as the four years following an election year. Following standard practice, the constant 0.01 is added to every district-level observation before taking logs in order to avoid completely unlit districts dropping out of the sample. The ability of nighttime lights to measure economic activity is established by Henderson et al. (2012) at the national level, while Hodler & Raschky (2014) and Michalopoulos & Papaioannou (2013) extend their scope to the subnational level.

Summary statistics for the relevant variables are shown in Table B.1 in the appendix to this chapter.

[2]The Guan do not represent the majority of the population in any district, which is why this group is neglected in the following analysis.

[3]Image and data processing by the Earth Observation Group, NOAA National Geophysical Data Center. DMSP data collected by US Air Force Weather Agency.

2.4 Empirical strategy

The empirical strategy of this study is composed of two elements. The first one exploits variation in the co-ethnicity status of the two groups that each are co-ethnic with the president and the ruling party at some point during the sample period. This variation is induced in districts where either the Ewe or the Asante represent the majority of the population by the change in government from the NDC to the NPP following the election in the year 2000. The corresponding two-way fixed effects regression model closely resembles the econometric approach of Burgess et al. (2015) and De Luca et al. (2016):

$$y_{it} = \alpha \cdot \textit{Co-Ethnic}_{it} + \beta \cdot \textit{Co-Ethnic}_{it} \cdot Term_t + \gamma \cdot Term_t + \delta_t + \mu_i + \epsilon_{it} \tag{2.2}$$

where the dependent variable are (log) nighttime lights in district i in year t and $\textit{Co-Ethnic}_{it}$ is a binary indicator variable that takes on the value 1 if a district's majority population is co-ethnic with the Ghanaian president during the respective time period and zero otherwise. A positive and statistically significant estimate of the coefficient α would be interpreted as evidence for ethnic favoritism. The marginal effect of co-ethnicity with the president can further be allowed to vary across presidential terms by interacting the co-ethnicity indicator with $Term_t$ dummies for three of the four terms. In line with the aforementioned timing of Ghanaian elections, the electoral victory of the NPP in 2000 alters the co-ethnicity status for the Asante- and Ewe-dominated districts only in 2001. Similarly, each of the presidential term dummies is switched to 1 only during the four years following the respective election year.

δ_t and μ_i represent year and district fixed effects respectively. While the former absorb country-wide economic shocks, as well as year-dependent measurement errors in the nighttime lights outcome variable, the latter control for time-invariant heterogeneity across districts. Errors are allowed to be clustered within districts over time.

Instead of incorporating the co-ethnicity indicator, it is also possible to directly consider the potential economic fortunes that accrue to districts where the Ewe or the Asante represent the majority of the population. Given that the binary indicator variables that reflect the presence of these population majorities are time-invariant, however, only their interaction effects with the presidential terms can be identified in the fixed-effects model.

The second element of the empirical strategy models variation in the districts' nighttime lights intensity as a result of the district-level interplay between close voting and the ethnic composition. While any indicators of ethnic composition are time-invariant within districts, the voting behavior varies across elections. Interacting the two hence allows identifying the marginal effects of ethnicity on the effect of close voting while maintaining the two-way fixed effects framework:

$$y_{it} = \alpha \cdot Competitiveness_{it} + \beta \cdot Competitiveness_{it} \cdot Ethnic_{ij} + \delta_t + \mu_i + \epsilon_{it} \tag{2.3}$$

where $Competitiveness_{it}$ is a measure of the electoral competitiveness of the districts that varies with the quadrennial elections. This study considers two such measures: the vote margin as an absolute measure of electoral tightness and the decile rank of the vote margin as a relative one. The vote margin is computed as the absolute difference between the vote shares of the NDC and the NPP, which are the only relevant parties. Hence, the vote margin changes within each district after every election, i.e. every four years. The vote margin is also the measure of choice in Banful (2011). Given that the latter reports a strong negative association between the vote margin and budget allocations of the central government to district authorities between 1994 and 2005, estimating Equation 2.3 without the interaction term can serve as a test of the ability of the nighttime lights proxy to pick up the effect of the politically induced spending. However, the vote margin is not necessarily comparable across all four elections in the sample. In order to take this potential pitfall into account, the vote margin of each district is assigned to its decile rank for each election separately.

The binary variable $Ethnic_{ij}$ then indicates whether ethnic group j represents the majority of the population in district i.

A district's urbanization rate and the average years of schooling of a district's population in the year 2000 are interacted with linear time trends. The interactions can be added to all aforementioned regression models to control for trends in these two variables that might affect a district's economic prosperity.

2.5 Results

2.5.1 Co-ethnicity and economic prosperity

Column 1 of Table 2.1 reports the result from estimating Equation 2.2 while omitting the interaction between the co-ethnicity indicator and the presidential terms, hence identifying only from the within-district variation in the co-ethnicity status. The estimate of the parameter α suggests that the effect of co-ethnicity with the Ghanaian president on nighttime lights is negative when averaging it over the entire sample period, but the estimate is statistically insignificant. This result is unaffected by controlling for district trends in urbanization and schooling (columns 2 and 3).

Column 1 of Table 2.2 shows results from repeating the estimation of Equation 2.2, but now allowing the effect of co-ethnicity to vary across the presidential terms. While omit-

Table 2.1: Co-ethnicity and economic prosperity 1992-2008

	(1)	(2)	(3)
Dep. Variable		Log Nighttime Lights	
Co-Ethnic	-0.057	-0.049	-0.056
	(0.064)	(0.062)	(0.054)
Observations	1,760	1,760	1,760
Within R^2	0.285	0.316	0.321
Number of clusters	110	110	110
District & Year FE	Yes	Yes	Yes
Urbanization rate 2000 x Trend	No	Yes	Yes
Years of schooling 2000 x Trend	No	No	Yes

Notes: Robust standard errors clustered at the district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

ting the interaction with the first presidential term (1993-1996), the parameter estimates for the three remaining terms reveal that the effect of co-ethnicity depends considerably on the government in power. The main effect of co-ethnicity is positive and significant, suggesting that co-ethnic districts exhibit on average a 34% ($\exp(0.291) - 1 = 0.3378$) higher nighttime lights intensity than districts that are not co-ethnic. However, the effect of co-ethnicity is considerably smaller in magnitude during the second presidential term ($\exp(0.291 - 0.218) - 1 = 0.073 = 7.3\%$). Then, after the NPP take-over of power at the beginning of the third presidential term and Asante-majority districts becoming co-ethnic with the president, the total effect of co-ethnicity is even negative and large in absolute size: Nighttime lights intensity is 21% lower ($\exp(0.291 - 0.526) - 1$) for co-ethnic districts. This relative disadvantage continues on a similar level during the second term of the NPP-led government from 2005 to 2008, which simultaneously represents the last presidential term in the sample. Taken together, these results suggest that ethnic favoritism appears to be weak in the Ghanaian context given that the association between co-ethnicity and economic prosperity becomes minuscule quickly after Ghana's return to democracy and then even turns negative after the country's first democratic transition of power.

Considering the changing fortunes of the Ewe and the Asante separately over the presidential terms reveals another interesting pattern: Column 1 of Table 2.3 shows that, as expected from the previous results, districts with a majority of Ewe among their population are relatively worse off during the second presidential term when the NDC is still in power. Notably, they continue to be worse off compared to the remainder of districts also when they are not co-ethnic with the president anymore: While the estimated interaction effect during the first NPP term is insignificant, it is negative, highly significant and large in absolute magnitude ($\exp(-0.378) - 1 = 31\%$ lower nighttime lights intensity) during the

Table 2.2: Co-ethnicity and economic prosperity by governments 1992-2008

	(1)	(2)	(3)
Dep. Variable	Log Nighttime Lights		
Co-Ethnic	0.291**	0.322***	0.260**
	(0.113)	(0.110)	(0.128)
Co-Ethnic x (1997-2000)	-0.218**	-0.230***	-0.207**
	(0.090)	(0.086)	(0.092)
Co-Ethnic x (2001-2004)	-0.526***	-0.561***	-0.476**
	(0.170)	(0.158)	(0.186)
Co-Ethnic x (2005-2008)	-0.538***	-0.573***	-0.476**
	(0.193)	(0.191)	(0.224)
1997-2000	0.754***	0.960***	1.042***
	(0.093)	(0.108)	(0.117)
2001-2004	0.786***	1.111***	1.235***
	(0.094)	(0.123)	(0.155)
2005-2008	1.117***	1.559***	1.731***
	(0.095)	(0.142)	(0.195)
Observations	1,760	1,760	1,760
Within R^2	0.292	0.324	0.326
Number of clusters	110	110	110
District & Year FE	Yes	Yes	Yes
Urbanization rate 2000 x Trend	No	Yes	Yes
Years of schooling 2000 x Trend	No	No	Yes

Notes: Robust standard errors clustered at the district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

second NPP term. This result holds when controlling for the urbanization and schooling trends (column 2) and when using the population share of the Ewe instead of the majority indicator (columns 3 and 4).

Conversely, districts where the majority of the population belongs to the Asante ethnicity are already worse off in the last presidential period before the NPP climbs to power, as reported in column 5. The magnitude of their disadvantage increases just slightly during the two subsequent terms when the Asante are co-ethnic with the president. While the estimated coefficient for the last NDC term (1997-2000) is insignificant ($p = 0.119$) when the population share of the Asante is used instead, it points in the same direction.

Table 2.3: Ewe and Asante districts' economic prosperity by governments 1992-2008

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable				Log Nighttime Lights				
Ewe Majority x (1997-2000)	-0.218**	-0.203**						
	(0.090)	(0.091)						
Ewe Majority x (2001-2004)	-0.152	-0.122						
	(0.109)	(0.116)						
Ewe Majority x (2005-2008)	-0.378***	-0.333**						
	(0.119)	(0.133)						
Share of Ewe x (1997-2000)			-0.003**	-0.003**				
			(0.001)	(0.001)				
Share of Ewe x (2001-2004)			-0.004**	-0.003*				
			(0.002)	(0.002)				
Share of Ewe x (2005-2008)			-0.007***	-0.006***				
			(0.002)	(0.002)				
Asante Majority x (1997-2000)					-0.354**	-0.340***		
					(0.135)	(0.129)		
Asante Majority x (2001-2004)					-0.398***	-0.370***		
					(0.131)	(0.112)		
Asante Majority x (2005-2008)					-0.409***	-0.367***		
					(0.119)	(0.117)		
Share of Asante x (1997-2000)							-0.004	-0.004
							(0.003)	(0.003)
Share of Asante x (2001-2004)							-0.007***	-0.006**
							(0.003)	(0.002)
Share of Asante x (2005-2008)							-0.006***	-0.005**
							(0.002)	(0.002)
Observations	1,760	1,760	1,760	1,760	1,760	1,760	1,760	1,760
Within R^2	0.289	0.324	0.294	0.327	0.294	0.328	0.296	0.327
Number of clusters	110	110	110	110	110	110	110	110
District & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Presidential term FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urbanization, Schooling 2000 x Trend	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Hence, in addition to the finding that co-ethnicity is not only not beneficial but actually disadvantageous in the Ghanaian context, the situation is even gloomier for the two politically invested groups of the Ewe and Asante: Even in periods when they are not co-ethnic with the incumbent president, they are worse off compared to the remainder of the country when their fierce political opponents are in power. However, this empirical pattern is entirely consistent with the predictions of the probabilistic voting theory: Both the NDC and the NPP can rely on the strong ideological bias of the Ewe and the Asante respectively. Therefore, neither party receives a payoff in terms of additional votes from redirecting economic resources neither towards the Ewe nor the Asante homelands. If the latter's inhabitants are tied to one of the two parties, they provide overwhelming support for it anyway, while they cannot be swayed by prospects of positive redistribution made by the other party. In fact, the ideological bias of both of the politically invested groups appears to be so strong that the respective party to which they are tied can actually redirect resources away from them.

2.5.2 Close voting, ethnicity and economic prosperity

The first column of Table 2.4 reports results from estimating Equation 2.3 using the vote margin between the NDC and the NPP as the measure of electoral competitiveness while not taking any measures of ethnicity into account yet. Sign and significance of the estimated coefficient confirm the finding of Banful (2011) that economic prosperity is negatively associated with the district-level vote margin of presidential elections in the sample period: A one-percentage point larger vote margin translates into $\exp(-0.007) - 1 = -0.007\%$ lower nighttime lights intensity on average. Hence, the tighter the electoral outcome, the stronger the surplus in economic prosperity in the following presidential term. This finding is robust to the inclusion of the district trends (column 2). Adding a squared term of the vote margin (column 3) does not indicate a relevant nonlinearity in the politico-economic relationship. Interacting the vote margin with the dummies for the presidential terms (column 4) further suggests that the effect of the former does not vary substantially across the latter; the weak statistical significance of the interactions eventually disappears with the inclusion of the district-level trends (column 5).

Table 2.4: Close elections and economic prosperity 1992-2008

	(1)	(2)	(3)	(4)	(5)
Dep. Variable			Log Nighttime Lights		
Vote Margin	-0.007***	-0.005***	-0.008**	-0.008***	-0.006**
	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)
Vote Margin x (1997-2000)				0.003*	0.002
				(0.002)	(0.002)
Vote Margin x (2001-2004)				0.003*	0.002
				(0.002)	(0.002)
Vote Margin x (2005-2008)				0.000	-0.002
				(0.002)	(0.003)
Vote Margin squared			0.000		
			(0.000)		
Observations	1,760	1,760	1,760	1,760	1,760
Within R^2	0.312	0.333	0.334	0.317	0.337
Number of clusters	110	110	110	110	110
District & Year FE	Yes	Yes	Yes	Yes	Yes
Presidential term FE	No	No	No	Yes	Yes
Urbanization, Schooling 2000 x Trend	No	Yes	Yes	No	Yes

Notes: Robust standard errors clustered at the district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

While column 1 of Table 2.5 reproduces the negative coefficient estimate of the vote margin from the previous table, column 2 then shows the result from interacting the vote margin variable with the indicator for the Akan-dominated districts. The main effect of the vote margin is now small in economic terms and only weakly significant, while the interaction is also negative, but highly significant and large in absolute magnitude. Jointly tested, the two coefficients are also highly significant ($p = 0.000$). Thus, the negative association between close voting and economic prosperity appears to be almost entirely driven by the close-voting behavior of the Akan-dominated districts. Controlling for the district trends further eliminates the statistical significance of the vote margin coefficient (column 3), while leaving the joint level of significance of the two coefficients unaffected ($p = 0.002$). The finding is unaltered when using the population share of the Akan instead of the majority indicator as a measure of the Akan presence (columns 4 and 5).

It is further possible to interact the Akan-vote margin term with the indicators for the presidential terms to test whether the association between the Akan voting and economic prosperity changes across terms. Column 6 shows that the effect of the Akan majority evaluated at each presidential term separately is then composed of three different coefficients: the coefficient of the interaction with the vote margin, the coefficient of the interaction with the term indicator, and the estimate of the triple interaction effect of the Akan indicator, the vote margin and the term indicator. Tested jointly, these coefficients are insignificant for every presidential term (1997-2000: $p = 0.825$, 2001-2004: $p = 0.603$, 2005-2008: $p = 0.93$).

Table 2.5: Close elections, Akan ethnicity and economic prosperity 1992-2008 I

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Variable	Log Nighttime Lights						
Vote Margin	-0.007***	-0.003*	-0.001	-0.002	0.001	-0.005**	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Vote Margin x Akan Majority		-0.010**	-0.011***			-0.012**	-0.011*
		(0.004)	(0.004)			(0.006)	(0.006)
Vote Margin x Share of Akan				-0.000**	-0.000**		
				(0.000)	(0.000)		
Vote Margin x (1997-2000)						0.003	0.002
						(0.002)	(0.002)
Vote Margin x (2001-2004)						0.005**	0.004*
						(0.002)	(0.002)
Vote Margin x (2005-2008)						0.000	0.000
						(0.003)	(0.003)
Akan Majority x (1997-2000)						-0.041	0.075
						(0.187)	(0.190)
Akan Majority x (2001-2004)						0.131	0.303
						(0.230)	(0.236)
Akan Majority x (2005-2008)						-0.019	0.211
						(0.247)	(0.266)
Vote Margin x Akan Majority x (1997-2000)						0.012**	0.010*
						(0.006)	(0.006)
Vote Margin x Akan Majority x (2001-2004)						-0.000	-0.003
						(0.008)	(0.008)
Vote Margin x Akan Majority x (2005-2008)						0.010	0.006
						(0.008)	(0.008)
Observations	1,760	1,760	1,760	1,760	1,760	1,760	1,760
Within R^2	0.312	0.326	0.349	0.322	0.346	0.347	0.372
Number of clusters	110	110	110	110	110	110	110
District & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Presidential term FE	No	No	No	No	No	Yes	Yes
Urbanization, Schooling 2000 x Trend	No	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Next, the indicator for the tightest decile of the vote margin instead of the vote margin itself is inserted in Equation 2.3. As a consequence of the tightening elections, the tightest decile of the 1992 election comprises vote margins up to 10.3 pp, for example, while this upper limit of the same decile shrinks to 5.8 pp for the 1996 election. As reported in column 1 of Table 2.6, districts that move into the tightest decile of the vote margin over time experience a significant increase in their nighttime lights intensity of $\exp(0.156) - 1 = 16.88\%$ on average. This positive association between close voting and economic prosperity is robust to the inclusion of the district-level trends; it hence corresponds to the negative effect of the vote margin reported in the previous table. Interacting the decile indicator with the Akan majority indicator also has a similar effect on the measure of the electoral tightness as before: Both without (column 3) and with district-level trends (column 4), the estimated coefficient of the decile indicator itself decreases in magnitude and loses significance, while the interaction in comparison is large in magnitude and significant. Irrespective of the district-level trends, the two estimates are jointly highly significant ($p = 0.000$ without district-levels trends, $p = 0.004$ with district-level trends included). However, the results differ in comparison to the ones obtained from using the vote margin as the tightness measure when interacting the Akan-decile interaction further with the presidential term indicators, as reported in column 5. The marginal effects of the Akan majority are now significantly different from zero for every term (1997-2000: $p = 0.005$, 2001-2004: $p = 0.002$, 2005-2008: $p = 0.077$). The sum of the three coefficients of interest is positive for each presidential term, with the largest marginal effect of close voting in an Akan-dominated district occurring during the term 2001-2004. This makes sense given this term follows the first election since the return to democracy in which the former dictator Rawlings could not run for office again, thereby potentially weakening the position of the NDC, with the NPP making use of this opportunity by swaying Akan voters into its camp with promises of future economic compensation.

Table 2.6: Close elections, Akan ethnicity and economic prosperity 1992-2008 II

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable			Log Nighttime Lights			
First Decile	0.156***	0.114**	0.052	0.033	0.289***	0.185
	(0.044)	(0.044)	(0.065)	(0.064)	(0.099)	(0.115)
First Decile x Akan Majority			0.186**	0.145*	0.444**	0.279
			(0.084)	(0.080)	(0.214)	(0.239)
First Decile x (1997-2000)					-0.381***	-0.335***
					(0.088)	(0.116)
First Decile x (2001-2004)					-0.438***	-0.257*
					(0.131)	(0.137)
First Decile x (2005-2008)					-0.215	-0.131
					(0.138)	(0.171)
Akan Majority x (1997-2000)					0.434**	0.445**
					(0.179)	(0.173)
Akan Majority x (2001-2004)					0.286	0.364*
					(0.205)	(0.193)
Akan Majority x (2005-2008)					0.405*	0.525***
					(0.211)	(0.198)
First Decile x Akan Majority x (1997-2000)					-0.363	-0.114
					(0.249)	(0.281)
First Decile x Akan Majority x (2001-2004)					-0.058	0.008
					(0.295)	(0.303)
First Decile x Akan Majority xx (2005-2008)					-0.531*	-0.370
					(0.304)	(0.324)
Observations	1,760	1,760	1,760	1,760	1,760	1,760
Within R^2	0.290	0.323	0.292	0.325	0.318	0.353
Number of clusters	110	110	110	110	110	110
District & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Presidential term FE	No	No	No	No	Yes	Yes
Urbanization, Schooling 2000 x Trend	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Conversely, interacting the vote margin variable with the indicators for the other ethnic groups does not any comparable effect. As reported in Table 2.7, the main effect of the vote margin remains negative and highly significant in the presence of the interactions. Only the interaction with the indicator for the Grusi group is significant, but its sign is positive (columns 7 and 8), while all other interaction effects are not statistically different from zero (columns 1-6). Essentially the same results are obtained when interchanging the vote margin with the indicator for the first decile, as shown in Table 2.8. While there is a weakly significant and negative interaction effect of the Gurma group and the decile indicator, it is worth mentioning that there is only one district in the sample where the Gurma represent the majority population which finds itself in the first vote margin decile only during one presidential term.

Table 2.7: Close elections, other ethnicities and economic prosperity 1992-2008 I

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable				Log Nighttime Lights				
Vote Margin	-0.007***	-0.005***	-0.007***	-0.005***	-0.006***	-0.005***	-0.007***	-0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Ga-Adangbe Majority	omitted	omitted						
	(.)	(.)						
Ga-Adangbe Majority x Vote Margin	0.011	0.010						
	(0.007)	(0.006)						
Mole Majority			omitted	omitted				
			(.)	(.)				
Mole Majority x Vote Margin			-0.001	0.002				
			(0.004)	(0.004)				
Gurma Majority					omitted	omitted		
					(.)	(.)		
Gurma Majority x Vote Margin					-0.003	-0.001		
					(0.003)	(0.004)		
Grusi Majority							omitted	omitted
							(.)	(.)
Grusi Majority x Vote Margin							0.006***	0.005***
							(0.001)	(0.001)
Observations	1,760	1,760	1,760	1,760	1,760	1,760	1,760	1,760
Within R^2	0.315	0.336	0.312	0.333	0.312	0.333	0.312	0.333
Number of clustered	110	110	110	110	110	110	110	110
District & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urbanization, Schooling 2000 x Trend	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.8: Close elections, other ethnicities and economic prosperity 1992-2008 II

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Variable				Log Nighttime Lights				
First Decile	0.156***	0.114**	0.176***	0.131***	0.156***	0.118***	0.158***	0.115**
	(0.044)	(0.044)	(0.044)	(0.044)	(0.045)	(0.044)	(0.045)	(0.045)
Ga-Adangbe Majority	omitted	omitted						
	(.)	(.)						
Ga-Adangbe Majority x First Decile	omitted	omitted						
	(.)	(.)						
Mole Majority			omitted	omitted				
			(.)	(.)				
Mole Majority x First Decile			-0.245	-0.208				
			(0.180)	(0.154)				
Gurma Majority					omitted	omitted		
					(.)	(.)		
Gurma Majority x First Decile					-0.017	-0.158*		
					(0.064)	(0.082)		
Grusi Majority							omitted	omitted
							(.)	(.)
Grusi Majority x First Decile							-0.079	-0.035
							(0.051)	(0.051)
Observations	1,760	1,760	1,760	1,760	1,760	1,760	1,760	1,760
Within R^2	0.290	0.323	0.292	0.324	0.290	0.323	0.291	0.323
Number of clustered	110	110	110	110	110	110	110	110
District & Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urbanization, Schooling 2000 x Trend	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the district level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.6 Discussion

A first observation from the previous sections that warrants discussion is the increasing competitiveness of the Ghanaian elections over time. Why have both the national and district-level vote margins of the 1992 election not been as close as those of the 2004 election? A potential explanation is that with Ghana's return to democracy in 1992, the Ghanaian society just began to accumulate what Persson & Tabellini (2009) label as democratic capital. The stock of democratic capital reflects the stability of democracy as a system of governance. This stability, in turn, is tested and demonstrated e.g. by democratic transitions of power, which happened in the year 2000 for the first time in the Ghanaian case. Prior to that, the stability of democracy might very well have seemed questionable with the former dictator holding the office of the president and his political aides potentially controlling the institutions of governance already since the beginning of Rawlings' military rule in 1981. But once democracy had proven itself to be dependable by another round of free elections in 1996 and a peaceful transition of power in 2000, the stock of democratic capital might have risen sufficiently to open up opportunities for both parties and voters for more intense democratic competition.

Next, the observed economic penalty for the two co-ethnic groups and the simultaneous favoritism of the close-voting Akan-dominated districts bring the potential reasons for this politico-economic equilibrium into focus. It appears particularly puzzling why the Akan do not grant more generous vote margins to the Asante-led NPP. Recall that ethnically, the Asante belong to the Akan family, which is why the two groups are more closely related to each other than each of them is to any other ethnic group in Ghana.

A historical perspective (Ward, 1966) on ethnic relations in Ghana is illuminative in this regard: Despite the absence of widespread inter-ethnic violence in post-independence Ghana and the common cultural and ethnic background of the Akan, the coexistence of the various Akan groups has not always been peaceful or amicable. Instead, in the course of the formation of the Ashanti Kingdom around 1700, the Asante had organized themselves into a militaristic society whose power surpassed that of the neighboring Akan tribes and kingdoms. The Asante gradually defeated the latter and made them tributaries to their empire during the 18th century. At the height of the Ashanti power, only the most northern territories, as well as the Akan-Fante tribes along the coast line and the Ewe in today's Volta region were not subjected to the empire's power.

At that time, the Europeans, mostly the British and Danes, possessed only a number of forts along the coast, which however played an important role in the transatlantic slave trade. Slaves were predominantly provided by the Ashanti empire in exchange for European weapons. The slaves, in turn, were either captives from the wars of the Asante against their neighbors or turned in as part of tributary agreements of the Asante with the already subjected Akan tribes. The Fante, whose homeland housed the European forts, first acted as intermediaries between the Asante and the Europeans (Hargrove, 2015), but then saw themselves increasingly confronted with the danger of an Asante invasion, too. This danger was only averted by a series of military British interventions against the Asante, which however lasted over almost the entire 19th century and only came to an end with the establishment of a British protectorate over Ashanti in 1897.

The transatlantic slave trade has been found to exert a persistent negative influence on inter-ethnic trust across Africa in the long-run (Nunn, 2011). In the context of the strained historical relation between the Asante and their proximate Akan neighbors, one potential explanation for why the Akan groups surrounding the present-day Ashanti region are reluctant to support the Asante unconditionally in the realm of politics is that the Akan do not have a strong ideological preference for an Asante-led government due to persistent distrust towards the latter. Instead, the Akan might prefer to constrain the political power of the Asante by signaling that while the NPP has the possibility to gain the votes of the Akan, it can never take them for granted, in particular not without providing benefits in return. Empirical investigations that could serve to back up this hypothesis are currently constrained by the lack of data on inter-ethnic trust in Ghana. While the Afrobarometer (2015) asks respondents how much they trust their own ethnic group and how much they trust other ethnicities in the respective country, its Ghanaian samples neither differentiate between the Akan and the Asante with regard to the identity of the respondents, nor do they differentiate between the other Ghanaian ethnicities with regard to the degree of trust that the respondents have towards them.

Besides the ethnic perspective, it should be kept in mind that the south of Ghana is generally more developed than the northern regions in terms of infrastructure, living standards and economic performance. Jedwab & Moradi (2016) find evidence that this economic gradient has largely been set up during Ghana's colonial period. Nowadays, it is potentially related to the household utilization of devices such as radio, television and the internet, which facilitate political information and campaigning. Evidence from Sierra Leone (Casey, 2015) suggests that radio broadcasting is important as a means to inform voters and to make them cross traditional ethnic-party lines. The economic backlog of the northern regions of Ghana could then translate into greater costs or difficulties for politicians to address the voters residing there. This in turn could disparately hinder the NPP in particular from campaigning effectively for votes in the north if the Northerners could not be exposed to new politicians and political programs, given that the NPP was the newcomer on the political landscape at the end of the Rawlings military rule. The sustained growth of the Ghanaian economy since the return to democracy clearly has the potential to diminish this potential economic obstacle, but at least regarding the sample

period used for this study, there is no evidence that the northern ethnicities have at some point been incorporated into the interplay of close voting and economic transfers yet.
Finally, the finding that close-voting Akan districts appear to be economically favored naturally raises the question how this favoritism materializes. While the methodological choice of nighttime lights as the outcome variable in the presented regressions reflects the approval of nighttime lights as a general proxy for economic prosperity, it clearly limits the possibilities to explore what the respective districts actually receive as a consequence of their contested votes. Official data publications on subnational conditions in terms of poverty, education and health, while scarce, might be able to provide at least suggestive evidence on the channels of favoritism in Ghana.

2.7 Conclusion

This study examines ethnic favoritism under the aspect of competitive, democratic elections in the ethnically diverse West African nation of Ghana. Its results suggest that the two co-ethnic ethnicities become economically disadvantaged in relative terms soon after the country's return to democracy. Further, there is a positive association between close voting and economic prosperity with regard to the politically unaffiliated group of the Akan. Taken together, while there is no evidence for ethnic favoritism between a political leader and her co-ethnics in Ghana, ethnic favoritism can exist outside of this particular constellation. Indeed, the Ghanaian case exemplifies the insight of the probabilistic voting theory that in a politico-ethnic setup where the votes of a politically unaffiliated group, here an ethnicity, are 'up for grabs', both ethnically tied parties engage in attempts to sway these votes into their respective camp by promises of economic redistribution.
The presented results stir a considerable number of new and further questions for research, some of which are expounded in Section 2.6. Additional data sources have to be brought up to conduct the necessary empirical investigations into both the roots and the functioning of ethnic favoritism in Ghana.
The more recent years of Ghana's democracy promise to provide even more insights into the role of ethnicity in the framework of political competition: Since 2008, co-ethnicity with the governing party has not translated into co-ethnicity with the president anymore. This opens up opportunities for new interactions between the ethnic and the political sphere that may dissolve (or reinforce) the patterns detected within the sample period of this study. Simultaneously, the competitiveness of Ghanaian elections has remained very high. Ghanaian voters hence remain reluctant towards granting unconditional support to their governments, thereby providing rich opportunities for future research.

Chapter 3

Applicability and Consistency of Nighttime Lights: A Systematic Evaluation

3.1 Introduction

Applications of nighttime lights to economics in general and to development economics in particular have become increasingly popular over the recent years (see Donaldson & Storeygard (2016) and Huang et al. (2014) for comprehensive reviews). Typically, nighttime lights are intended to proxy economic variables in settings where the data sources for the latter are either unavailable and/or unreliable. Therefore, the attractiveness of nighttime lights is undoubtedly born out of necessity. Consecutive and global data on nighttime lights are currently available in yearly intervals for the time period from 1992 to 2013. This period is marked both by the urge for and by the prevalent absence of reliable, consistent, and sufficiently frequent quantitative measurements of economic advancement across the world. Official statistics from developing countries are particularly rarely available below the regional level and if they are, they are often the result of small-area estimations which are not only data-intensive and therefore infrequent, but also not without caveats (Deaton & Tarozzi, 2009). Hence, one cannot evade the idea that if nighttime lights contained any relevant information that could supplement and even extent the existing data, then they should indeed be made use of as far as possible.

The wide range of applications raises the question what nighttime lights then actually *are* if they apparently represent a strong and universal proxy variable. A rather blunt, but not necessarily incorrect response would suggest that nighttime lights are whatever variable the specific research question needs them to be. However, their strong proxy properties are certainly not accepted *prima facie*: Fortunately, it is considered good practice to test the applicability of nighttime lights to the respective research context, mostly by correlating them against the variables to be proxied over those time periods and levels of aggregation for which data are at least partially available. But because the diversity of applications has become so vast, the respective tests of applicability, while valid on their own, are not directly comparable with each other, making it difficult to derive general conclusions about applicability and consistency of nighttime lights from them.

For example, Henderson et al. (2012) correlate national GDP growth rates against growth rates of nighttime lights in a cross-country panel from 1992 to 2008. Chen & Nordhaus (2011) and Chen & Nordhaus (2015) also consider GDP and GDP growth at the national level, while adding very disaggregated information from the level of grid cells. Hodler & Raschky (2014) test their panel dataset of regional nighttime lights with the help of regional GDP data collected by Gennaioli et al. (2013), but the set of countries for which the regional GDP is available is both smaller than the set of countries that Hodler & Raschky (2014) consider in their full analysis and than the set of countries used by Henderson et al. (2012). Bhandari & Roychowdhury (2011) also apply nighttime lights at the sub-national level, but their correlations of nighttime lights and GDP are estimated at the more disaggregated level of districts. Further, they use a single cross-section of data and restrict their sample to the country of India. Michalopoulos & Papaioannou (2013), in turn, focus on the sub-national level of sub-Saharan African countries and correlate their cross-sectional nighttime lights data against a composite wealth index constructed at the level of enumeration areas of the Demographic and Health Surveys (DHS) across four countries.

On top of all these different levels of aggregations, time frames, outcomes and countries, furthermore also the exact computation of the nighttime lights measure is not the same across applications, but mostly varies between using lights per pixel versus lights per capita across a delimited territory. This ever-increasing diversity of applications does not raise any doubts about the validity of nighttime lights in all these specific contexts, but at the same time, it impairs the derivation of general statements about the consistency of nighttime lights applications.

This paper takes a systematic approach to evaluate the applicability and consistency of nighttime lights in the development nexus. By "systematic", I mean that the approach attempts to hold as many elements of the empirical framework fixed as possible while switching single parameters on and off one after another. The elements that I hold fixed permanently in the following are the sources of the data, the time frame, the country identity, and the variables that the nighttime lights are supposed to correlate with. Doing so allows me to examine the applicability and consistency of nighttime lights along two important dimensions: Firstly, the level of spatial aggregation can be shifted between the regional and the more disaggregated district level. Secondly, the correlations can be estimated either by pooling all observations or by exploiting the panel structure of my data. While the four resulting combinations are by far not exhaustive of the application potential of nighttime lights, they still provide some relatively clean-cut evidence on the questions whether nighttime lights can be discretionarily utilized in different spatial frameworks without loss of consistency and whether nighttime lights correlate as well with variables in levels as with their changes over time.

The results show that nighttime lights significantly correlate with two consistently measures variables of interest from the development context across all four combinations of applications. The correlations appear to be stronger in terms of the estimated slope at the higher level of spatial aggregation, while they are equally stronger in the cross-sectional applications compared to the applications which exploit the panel structure.

Section 3.2 presents the data, while Section 3.3 outlines the empirical strategy of the evaluation. Section 3.4 presents the estimation results. Section 3.5 concludes and provides some suggestions for applying night lights as a proxy variable in the context of development.

3.2 Data

3.2.1 Nighttime lights data

The utilized nighttime lights dataset is the Version 4 DMSP-OLS Nighttime Lights Time Series provided by the National Oceanic and Atmospheric Administration (NOAA).[1] It is identical to the one commonly used in the literature and covers the period from 1992 to 2013. A total of six different satellites were involved in the recording process. Due to the dissemination of the dataset, detailed descriptions of this process and of the general properties of the data are omitted at this point. In brief, the yearly data is composed of pixels containing information on the average visible, stable lights. Each pixel is assigned a value of nighttime lights intensity between 0 and 63, with 63 representing the maximum level of luminosity that the satellites can record. The NOAA data are therefore affected by the technical problem of both top- and bottom-coding: In economically advanced and densely populated areas, the satellite data do not show any variation above the peak value of 63, while they are also invariant across areas that do not exhibit any stable lights.

The calculations of the mean value of nighttime lights within a georeferenced area are performed by the software QGIS (2017). While there are no self-evident provisions whether nighttime lights per geographic area, nighttime lights per capita, or some other form of population-adjusted nighttime lights variable are the superior transformation of choice,

[1] Image and data processing by NOAA's National Geophysical Data Center. DMSP data collected by US Air Force Weather Agency.

the focus is laid on the logarithm of nighttime lights per geographic area in the following, as this measure is the most universal one in the sense that it does not depend on the availability of additional data such as population figures. Instead, the nighttime lights per geographic area are simply the result of the software counting the total number of pixels that are comprised within an area's boundary, summing up the luminosity values of the pixels within the same area, and dividing the sum by the total number of pixels. Before taking logs, the value 0.01 is added to each observation in order to avoid observations dropping out of the sample because their average luminosity is zero, i.e. they are completely unlit.

3.2.2 Development data

IPUMS-International (2017) provides comprehensive extracts from census micro data for various developing countries. These extracts qualify for my analysis because firstly, they contain multiple variables which are of general interest in the context of development economics. Secondly, the data cover two subsequent census waves for several countries, allowing it to either pool all the observations or to arrange the data in a panel format and exploit only within-group changes over time.[2] Thirdly, they contain georeferenced information regarding both the first and the second sub-national level of the administrative location of every observation. I refer to administrative divisions at the first sub-national level as 'regions' and to administrative divisions at the second sub-national level as 'districts'. The census extracts are designed such that by using the specified weights, the aggregation of the individual- and household-level information to the respective administrative level yields representative averages. The location information is furthermore spatially harmonized over time, meaning that within each country, the number of regions and districts is hypothetically held constant over time, which greatly facilitates the arrangement of the data in panel format.

Fortunately, IPUMS-International (2017) also provides georeferenced maps for each country with harmonized external and internal administrative boundaries over time. I utilize the maps to calculate the mean of the nighttime lights within each administrative division, which can then be easily matched with the census information. In addition, I determine the latitudes and longitudes of the centroids of every administrative unit. A complete list of countries included in the sample together with summary statistics for the variables can be found in the appendix to this chapter.[3]

I select two variables from the census extracts data that I correlate against nighttime

[2] The first census year in the sample is 1998, the last one is 2011.

[3] Countries in the sample: Bangladesh, Cambodia, Ecuador, Ghana, Malawi, Mali, Zambia. Number of regions: 186. Number of districts: 1482.

lights in my application: the average rate of electrification of households and the average years of schooling of the population within each district or region. This choice is partly determined by the fact that among all variables, these two are most frequently available for countries for which there is also more than one wave of census data. Nevertheless, both variables can also be assumed to be quite comparable across countries and they are both frequently incorporated in multidimensional poverty indices, hinting at their relevance in the development context. Finally, electrification and schooling are interesting to correlate against nighttime lights because with regard to electrification, it is easy to presume a strong technical relationship with the local area's luminosity, while with regard to schooling, there is less reason to expect a strong relationship a priori. Additionally, I compute the total population of each administrative unit in order to then obtain a measure of population density by dividing the population figure by the number of pixels comprised by the respective unit's boundaries.

While variables that are directly related to income would also be very interesting to correlate against nighttime lights, such information is typically not recorded by census statistics. The dataset on regional GDP collected by Gennaioli et al. (2013) is not publicly available and even if it was, it would actually deprive my analysis of the possibility of shifting between different levels of spatial aggregation. Small-area poverty estimates obtained by applying the methodology by Elbers et al. (2003) are available for a number of developing countries and have been correlated against nighttime lights by Chen (2015). However, they are not directly comparable to GDP and they cannot be simply aggregated to higher spatial levels. Their accuracy has further been questioned by Deaton & Tarozzi (2009).

3.3 Empirical strategy

3.3.1 Estimation setup

Applicability and consistency of nighttime lights are to be evaluated based on their correlations with the two development outcomes at the two different levels of spatial aggregation and within the cross-sectional, as well as the longitudinal framework. Consequently, the basic setup consists of regressing the outcome y on the logarithm of nighttime lights $ln(Lights)$ in order to estimate the parameter β_1. The logarithm of the population density $ln(PopulationDensity)$ is added as the single control variable that varies at the spatial, as well as at the temporal level. However, as Bickenbach et al. (2016) point out, control variables should actually not strongly affect the correlations because the need for control variables would undermine the universal applicability of nighttime lights. Year fixed effects in every specification control for the effect of changing satellites recording the nighttime lights. When pooling the observations from all time periods, longitude and latitude of each district's or region's centroid are included in order to reduce any spatial autocorrelation, resulting in the following specification at the regional level which can be estimated by OLS:

$$y_j = \beta_0 + \beta_1 ln(Lights)_j + \beta_2 ln(PopulationDensity)_j + \gamma_1 X_j + \gamma_2 Y_j + \delta Year_t + \epsilon_j \quad (3.1)$$

When estimating the relationship at the district level, the error terms are allowed to be correlated at the regional level, advising the clustering of the standard errors at the same level:

$$y_i = \beta_0 + \beta_1 ln(Lights)_i + \beta_2 ln(PopulationDensity)_i + \gamma_1 X_i + \gamma_2 Y_i + \delta Year_t + \epsilon_{ij} \quad (3.2)$$

In order to exploit the panel structure of the data, group fixed effects absorb all time-invariant heterogeneity between the observations, estimating β_1 only from within-region variation:

$$y_{jt} = \beta_0 + \beta_1 ln(Lights)_{jt} + \beta_2 ln(PopulationDensity)_{jt} + \alpha Region_j + \delta Year_t + \epsilon_{jt} \quad (3.3)$$

The corresponding model at the district level reads:

$$y_{it} = \beta_0 + \beta_1 ln(Lights)_{it} + \beta_2 ln(PopulationDensity)_{it} + \alpha District_i + \delta Year_t + \epsilon_{it} \quad (3.4)$$

The errors are allowed to be serially correlated within groups, which suggests clustering the standard errors within each region respectively district over time.

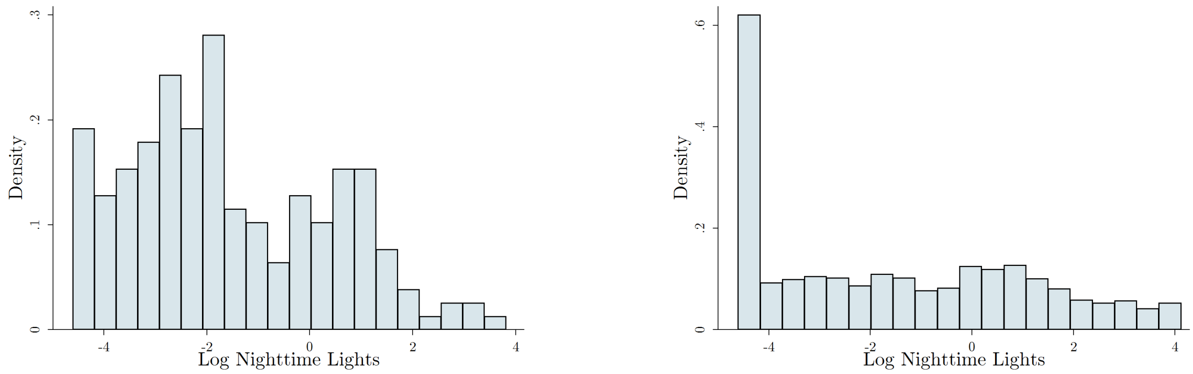
3.3.2 Econometric considerations

Within the outlined estimation setup, it is worthwhile to consider econometric challenges which have the potential to affect the consistency of the estimation of β_1, the parameter of the nighttime lights measure, and its inference. For example, a factor exerting a downward bias on the estimated standard errors of β_1 increases the danger of type I errors, i.e. of incorrectly rejecting the null of no significant relationship between the development outcomes and nighttime lights, while a downward bias on the parameter estimate may result in type II errors, i.e. of incorrectly inferring no significant relationship when it is actually there. It is particularly useful to consider these challenges along the two dimensions of district vs. region and pooled vs. panel, as the estimation setup intends to make shifts along one dimension while keeping the other one fixed.
First, Henderson et al. (2012) assume that both the outcome variable, in their case GDP growth, and the nighttime lights variable are each measured with 'classical' measurement

errors, with the errors being uncorrelated between the two variables. It as a standard result from econometric theory (Pischke, 2007) that measurement error in the dependent variable does not affect the consistency of the OLS estimates of β_1, but only inflates the estimated standard errors. Measurement error in the explanatory variable, however, causes an 'attenuation' bias in the estimates of β_1, i.e. biases them towards zero. Thus, under the assumption of measurement errors in both the development outcomes and the nighttime lights, all conventional OLS estimates of β_1 underestimate the true relationship between the outcomes and the nighttime lights. Furthermore, it is well-established that the attenuation bias from the measurement error in the explanatory variable is aggravated if β_1 is estimated within the framework of a fixed-effects model such as Equation 3.3 and Equation 3.4. This results from the within-transformation wiping out a lot of variation from the data, leaving only deviations from the group means to estimate the parameter, thereby diminishing the signal-to-noise ratio and amplifying the effect of the measurement error as part of the noise. Consequently, under the assumption of measurement errors in both the development outcomes and the nighttime lights, all fixed-effect estimates of β_1 will be smaller in terms of absolute size than the estimates of β_1 obtained from the pooled regressions.

Second, the disaggregated application of nighttime lights introduces an econometric problem that is not considered by Henderson et al. (2012) and Chen & Nordhaus (2011) because it is not relevant at the national level that they consider: Breaking down the spatial unit of observation into ever smaller parts increases the number of observations that are either top-coded because all pixels within them have a luminosity value of 63 or bottom-coded because all pixels are completely unlit. For example, on the national level, the mean nighttime lights intensity never takes the value 0 nor the value 63 (excluding city states), but disaggregating a developing country into its metropolitan and remote areas respectively will result in some fully and some unlit observations. This is equivalent to the nighttime lights measure being both left- and right-censored at the sub-national level, with the frequency of censoring increasing in the level of disaggregation. Indeed, while only five regions in the sample are completely unlit and no region is fully lit, 24.90% of the district-level observations have a mean luminosity of zero and 0.13% are top-coded. Figure 3.1 displays histograms of the distributions of the logarithm of nighttime lights at (a) the district and (b) the regional level.

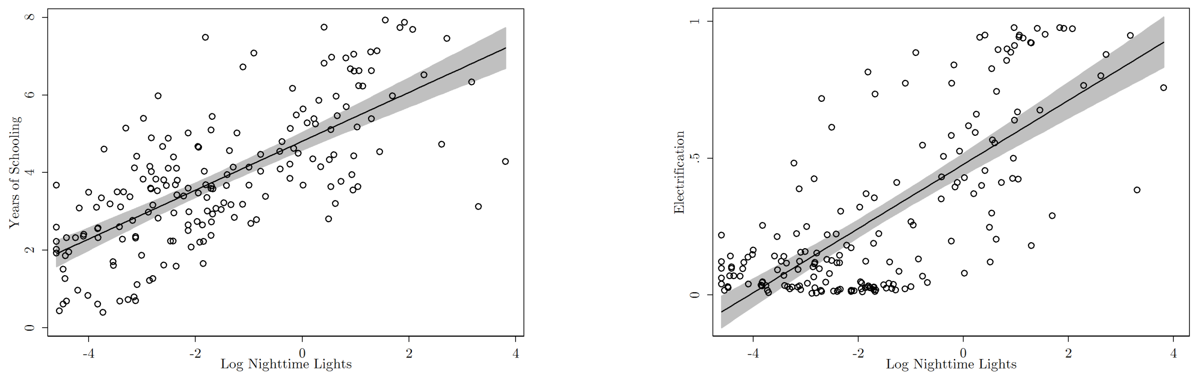
Figure 3.1: Histograms of nighttime lights at two levels of aggregation



(a) Regions (b) Districts

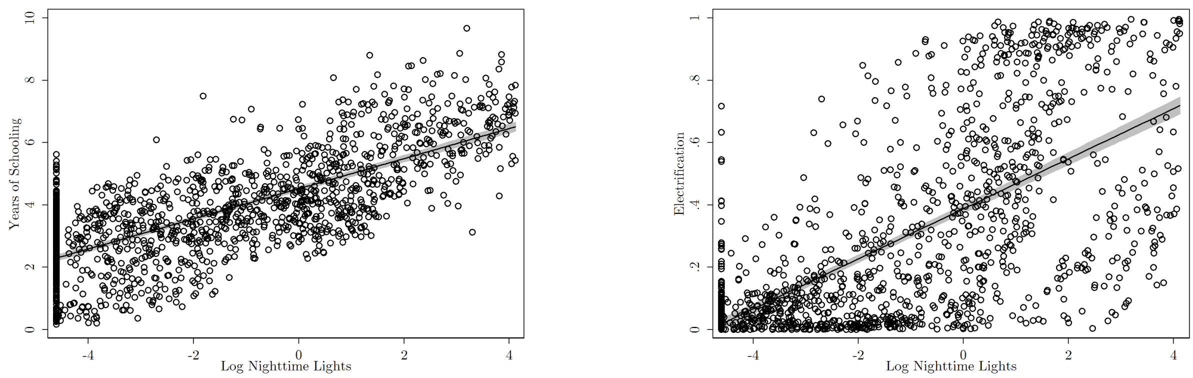
While the econometric literature has dealt extensively with the effects of a censored dependent variables, the effects of censoring in the explanatory variable have received less attention. Only relatively recently, Rigobon & Stoker (2007) and Rigobon & Stoker (2009) have shown that left-censoring in the regressor causes attenuation bias if, graphically speaking, the left-censored observations have their center of mass above the regression line, thereby 'pulling' the regression line upwards at its lower end, which is equivalent to attenuating its slope. Figure 3.3 shows scatter plots of the two development outcomes against the (log) nighttime lights in levels using the district-level observations. At the lower end of the nighttime lights spectrum, a large number of censored observations indeed piles up around the estimated best linear fit line, with the majority being located above the line. While the plots are noisier when the regional observations are used instead (Figure 3.2), there is much less indication of censored observations piling up at either the low- or the high-end of the nighttime lights spectrum. Given the strong prevalence of left-censoring in the district sample of nighttime lights, regressions estimating the nighttime lights parameter β_1 at the district level are expected to suffer more from attenuation bias than regressions at the regional level.

Figure 3.2: Scatter plots in levels at the regional level



(a) Years of Schooling (b) Electrification

Figure 3.3: Scatter plots in levels at the district level



(a) Years of Schooling (b) Electrification

3.4 Results

3.4.1 Results at the regional level

Table 3.1 reports results from regressing the schooling measure on nighttime lights and the expanding set of controls on the regional level. The raw correlation (column 1) remain positive and strongly significant with the addition of the population density measure (column 2), longitudes and latitudes (column 3), and the year fixed effects (column 4). Based on the most exhaustive specification, a 1% increase in nighttime lights intensity is associated with 0.005 more years of schooling. The estimates are naturally smaller (electrification is bound between 0 and 1), but similarly robust when the regional electrification rate is used as the dependent variable (Table 3.2), suggesting that 1% increase in nighttime lights intensity is associated with a 0.001 higher electrification rate.

Table 3.1: Pooled regressions of nighttime lights and schooling at the regional level

	(1)	(2)	(3)	(4)
Dep. Variable	Years of Schooling			
Log Nighttime Lights	0.630***	0.816***	0.791***	0.495***
	(0.055)	(0.056)	(0.083)	(0.079)
Log Population Density		-0.382***	-0.343***	-0.135*
		(0.083)	(0.098)	(0.082)
Observations	186	186	186	186
R-squared	0.498	0.571	0.630	0.749
Longitude/Latitude	No	No	Yes	Yes
Year FE	No	No	No	Yes

Notes: Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.2: Pooled regressions of nighttime lights and electrification at the regional level

	(1)	(2)	(3)	(4)
Dep. Variable	Electrification			
Log Nighttime Lights	0.117***	0.164***	0.137***	0.111***
	(0.008)	(0.011)	(0.014)	(0.016)
Log Population Density		-0.097***	-0.070***	-0.052***
		(0.013)	(0.015)	(0.016)
Observations	186	186	186	186
R-squared	0.530	0.677	0.770	0.809
Longitude/Latitude	No	No	Yes	Yes
Year FE	No	No	No	Yes

Notes: Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

When only accounting for region fixed effects in table 3.3, the raw correlation (column 1) between schooling and nighttime lights is actually estimated to be of larger magnitude than in the cross-sectional specification. However, after adding the population density control (column 2) and accounting for the year fixed effects, the parameter is quantitatively smaller (column 3), but still highly significant in statistical terms. With regard to electrification (Table 3.4), the estimated coefficients also drops in size by about two-thirds once all controls are included, but remain significant (p-value in column 3: 0.011).

Table 3.3: Panel regressions of nighttime lights and schooling at the regional level

	(1)	(2)	(3)
Dep. Variable		Years of Schooling	
Log Nighttime Lights	1.095***	0.811***	0.311***
	(0.078)	(0.067)	(0.077)
Log Population Density		1.283***	0.108
		(0.155)	(0.392)
Observations	186	186	186
R-squared	0.698	0.785	0.906
Number of Clusters	93	93	93
Region FE	Yes	Yes	Yes
Year FE	No	No	Yes

Notes: Standard errors clustered within regions in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Panel regressions of nighttime lights and electrification at the regional level

	(1)	(2)	(3)
Dep. Variable		Electrification	
Log Nighttime Lights	0.097***	0.049***	0.032**
	(0.012)	(0.008)	(0.012)
Log Population Density		0.215***	0.155***
		(0.029)	(0.053)
Observations	186	186	186
R-squared	0.377	0.545	0.761
Number of Clusters	93	93	93
Region FE	Yes	Yes	Yes
Year FE	No	No	Yes

Notes: Standard errors clustered within regions in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

3.4.2 Results at the district level

Table 3.5 reports results from district-level regressions of the schooling measure on nighttime lights and the various controls. All estimates of the nighttime lights parameter are positive and highly significant, suggesting that about 0.004 more years of schooling are associated with a 1% stronger nighttime lights intensity. Notably, variation in nighttime lights alone explains more than 50% of the district-level variation in schooling (column 1). The estimated coefficients are also strongly significant with regard to the electrification measure. Again, without any controls, nighttime lights already explain 48% of the variation in the outcome measure (column 1).

Table 3.5: Pooled regressions of nighttime lights and schooling at the district level

	(1)	(2)	(3)	(4)
Dep. Variable	Years of Schooling			
Log Nighttime Lights	0.483***	0.510***	0.463***	0.372***
	(0.034)	(0.042)	(0.041)	(0.035)
Log Population Density		-0.064	0.011	0.079
		(0.069)	(0.051)	(0.048)
Observations	1,482	1,482	1,482	1,482
R-squared	0.545	0.547	0.618	0.712
Number of Clusters	93	93	93	93
Longitude/Latitude	No	No	Yes	Yes
Year FE	No	No	No	Yes

Notes: Standard errors clustered within regions in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.6: Pooled regressions of nighttime lights and electrification at the district level

	(1)	(2)	(3)	(4)
Dep. Variable	Electrification			
Log Nighttime Lights	0.080***	0.103***	0.071***	0.060***
	(0.011)	(0.012)	(0.007)	(0.007)
Log Population Density		-0.056***	-0.009	-0.006
		(0.016)	(0.008)	(0.009)
Observations	1,482	1,482	1,482	1,482
R-squared	0.482	0.541	0.679	0.727
Number of Clusters	93	93	93	93
Longitude/Latitude	No	No	Yes	Yes
Year FE	No	No	No	Yes

Notes: Standard errors clustered within regions in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

By applying the within-estimator to Equation 3.4, all time-invariant heterogeneity between districts is wiped out. Nevertheless, the estimation retains a large, positive and highly significant correlation between schooling and nighttime lights (Table 3.7, column 1). Accounting for year-specific shocks sharply reduces the size of the estimated parameter (column 3), while leaving its statistical significance unaffected. The same pattern emerges with regard to the association of electrification with nighttime lights (Table 3.8): The correlations decrease in terms of magnitude, but remain positive and significant.

Table 3.7: Panel regressions of nighttime lights and schooling at the district level

	(1)	(2)	(3)
Dep. Variable		Years of Schooling	
Log Nighttime Lights	0.590***	0.445***	0.109***
	(0.039)	(0.034)	(0.023)
Log Population Density		1.492***	-0.378**
		(0.142)	(0.157)
Observations	1,482	1,482	1,482
R-squared	0.327	0.518	0.865
Number of Clusters	741	741	741
District FE	Yes	Yes	Yes
Year FE	No	No	Yes

Notes: Standard errors clustered within districts in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.8: Panel regressions of nighttime lights and electrification at the district level

	(1)	(2)	(3)
Dep. Variable		Electrification	
Log Nighttime Lights	0.070***	0.054***	0.026***
	(0.006)	(0.006)	(0.006)
Log Population Density		0.171***	0.014
		(0.025)	(0.046)
Observations	1,482	1,482	1,482
R-squared	0.222	0.342	0.617
Number of Clusters	741	741	741
District FE	Yes	Yes	Yes
Year FE	No	No	Yes

Notes: Standard errors clustered within districts in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

3.5 Conclusion

The intention of this paper is to perform a systematic evaluation of the applicability and consistency of nighttime lights. My findings show that firstly, nighttime lights correlate strongly with both the electrification rate of households and the average years of schooling in a sample of developing countries. These correlations confirm that, within my limited framework, the applicability of nighttime lights as a proxy clearly goes beyond variables which feature a somewhat direct relationship to the emittance of lights such as electrification or industrial production, as the association of nighttime lights and schooling is equally strong in terms of statistical significance and also linear in shape.

Secondly, shifting the level of spatial aggregation from the regional to the district level does not impair the ability of nighttime lights to correlate with the two development outcomes. While the district level is far from exhaustive with regard to the possibilities of disaggregating nighttime lights (or other remote sensing data), the district is already a relatively small administrative entity, suggesting that nighttime lights bear the potential to be usefully applied also within individual countries if the number of disaggregated administrative units is large enough in order to provide sufficient spatial variation.

Thirdly, nighttime lights remain strongly correlated with electrification and schooling even if a large amount of variation is discarded by estimating fixed-effects models and identifying the parameter of nighttime lights only from within-region or within-district variation. This suggests that nighttime lights are not just a stationary spatial feature, but that they co-evolve with the underlying economic, demographic and social patterns. However, the temporal gap between two observations of the same spatial unit is quite large in my sample due to the fact that population censuses in particular in developing countries are rarely performed more than once per decade. It would therefore be worthwhile to attempt to construct a similarly comparable dataset in which however the observations are recorded at shorter frequencies in order to assess whether nighttime lights pick up year-to-year variation in the outcomes as well as variation accumulated over several years.

Taken as a whole, the results of this evaluation suggest that nighttime lights maintain their consistency as a proxy variable when varying the parameters of the application systematically along the spatial and temporal dimension respectively. The assessment that nighttime lights correlate strongly and quite unconditionally with the years of schooling as a measure of educational attainment might be interpreted as particularly encouraging in the context of developing countries where disaggregated and consecutive educational data are at least as rare as GDP-related data. The correlation between nighttime lights and schooling further calls for replacing schooling with rather quality- and skills-oriented measures of human capital which Hanushek & Woessmann (2012a) and Hanushek & Woessmann (2012b) deem as much more relevant and informative than the mere length

of education.
The fact that the estimated slope parameters are not equal across the four different regression frameworks should not immediately be interpreted as a warning sign that nighttime lights as a proxy do not exhibit a constant quantitative association with the variables they are supposed to reflect. I point to a series of econometric specifics that should be kept in mind when nighttime lights act as the explanatory variable because certain properties of the nighttime lights data can result in downward biases in the parameter estimates. While the two specifics of classical measurement error and censoring in the explanatory variable are unlikely to cover all econometric challenges brought up by the chosen empirical framework, the size of the parameter estimates behaves exactly as the former imply it along the lines of the latter.
With regard to future research in the domain of remote sensing and nighttime lights in particular, it will be illuminating to examine the properties of the VIIRS-DNB[4] series in a framework similar to the one suggested in this paper once multiple yearly composites of stable lights will be available. According to Elvidge et al. (2013), the VIIRS-DNB series is superior to the DMSP-OLS data in terms of spatial resolution and detection limits. Therefore, it has the potential not only to tap into even lower levels of disaggregation, but also to alleviate the top- and bottom-coding issues at least for upcoming applications of nighttime lights.

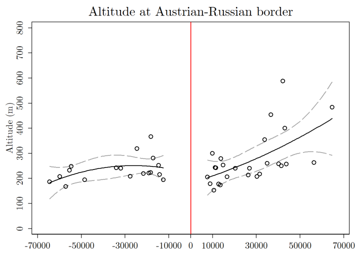
[4]Visible Infrared Imaging Radiometer Suite-Day/Night Band, Earth Observation Group, NOAA National Geophysical Data Center

Appendix A

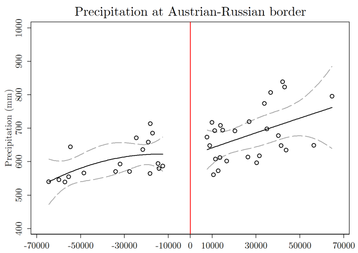
Appendix to Chapter 1

A.1 Figures

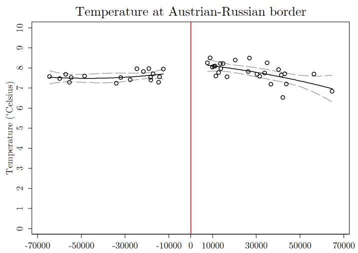
Figure A.1: Discontinuities in geographic characteristics at Austrian-Russian border



(a) Altitude



(b) Precipitation



(c) Temperature

A.2 Tables

Table A.1: Summary statistics at Austrian-Russian border

Variables	(1) N	(2) Mean	(3) SD	(4) Min	(5) Max
Population 1810	187	1413	1923.84	167	23612
Austrian-Russian Distance (m) 1910-1932	44	29,517	16,972	7,660	64,730
Austrian-Russian Distance (m) 1960	59	31,352	18,327	2,766	64,823
City Dummy 1910-1932	44	0.0455	0.211	0	1
City Dummy 1960	59	0.203	0.406	0	1
Average Altitude (m)	44	257.2	86.67	153.0	587.9
Average Temperature (Degree Celsius)	44	7.718	0.412	6.526	8.506
Average Precipitation (mm)	44	647.4	78.63	539.4	839.0
Average Caloric Yield	44	1,927	167.0	1,447	2,129
Optimal Caloric Yield	44	9,755	435.6	8,693	10,499
Share of Cropland 1800	44	0.501	0.110	0.293	0.679
Share of Cropland 1900	44	0.683	0.134	0.387	0.898
Primary Enrollment 1911	44	0.489	0.260	0.0880	0.862
Primary Enrollment 1921	44	0.720	0.168	0.409	1.006
Primary Enrollment 1926	44	0.692	0.111	0.415	1.103
Primary Enrollment 1932	44	0.994	0.0482	0.850	1.101
Primary Enrollment 1961	59	1.055	0.027	0.984	1.142
Share of Literates 1921	44	0.581	0.123	0.340	0.822
Share of Literates 1931	44	0.561	0.0799	0.402	0.806
Share of Literates 1960	59	0.954	0.0265	0.881	0.991
Schools Per 1,000 School-Aged 1911	44	3.026	1.098	1.028	5.644
Schools Per 1,000 School-Aged 1921	44	5.124	1.408	1.819	7.953
Schools Per 1,000 School-Aged 1926	44	5.113	1.185	1.584	7.397
Schools Per 1,000 School-Aged 1932	44	5.059	1.164	1.342	7.126
Schools Per 1,000 School-Aged 1961	59	5.71	2.38	1.137	12.01
Classes Per 1,000 School-Aged 1911	44	6.580	3.757	1.069	13.78
Classes Per 1,000 School-Aged 1921	44	9.175	2.352	5.872	15.62
Classes Per 1,000 School-Aged 1926	44	15.66	3.864	7.924	21.82
Teachers Per 1,000 School-Aged 1911	44	7.152	4.137	1.028	16.55
Teachers Per 1,000 School-Aged 1921	44	12.09	3.795	6.387	25.36
Teachers Per 1,000 School-Aged 1926	44	14.69	3.483	9.754	24.77

Table A.2: Summary statistics at Prussian-Russian border

Variable	(1) N	(2) Mean	(3) SD	(4) Min	(5) Max
Population 1810	165	1236.71	1566.91	126	16114
Prussian-Russian Distance (m) 1910-1932	54	32,154	18,902	6,521	64,629
Prussian-Russian Distance (m) 1960	53	34,677	18,669	3,728	64,741
City Dummy 1910-1932	54	0.0741	0.264	0	1
City Dummy 1960	53	0.208	0.409	0	1
Average Altitude (m)	54	107.0	32.23	28.27	182.6
Average Temperature (Degree Celsius)	54	7.917	0.421	7.023	8.474
Average Precipitation (mm)	54	541.3	21.01	514.4	607.4
Average Caloric Yield	54	2,023	131.2	1,686	2,213
Optimal Caloric Yield	54	9,802	340.6	8,992	10,466
Share of Cropland 1800	54	0.631	0.0448	0.517	0.695
Share of Cropland 1900	54	0.834	0.0590	0.683	0.918
Primary Enrollment 1911	54	0.693	0.395	0.0906	0.996
Primary Enrollment 1921	54	0.785	0.188	0.376	1.008
Primary Enrollment 1926	54	0.744	0.0816	0.506	0.922
Primary Enrollment 1932	54	1.010	0.0634	0.901	1.138
Primary Enrollment 1961	53	1.065	0.045	0.904	1.250
Share of Literates 1921	54	0.669	0.124	0.459	0.844
Share of Literates 1931	54	0.654	0.106	0.451	0.808
Share of Literates 1960	53	0.964	0.028	0.912	0.998
Schools Per 1,000 School-Aged 1911	54	5.299	3.053	0.795	11.43
Schools Per 1,000 School-Aged 1921	54	7.078	2.184	1.051	13.70
Schools Per 1,000 School-Aged 1926	54	7.387	2.392	0.820	11.88
Schools Per 1,000 School-Aged 1932	54	7.242	2.454	1.181	12.42
Schools Per 1,000 School-Aged 1961	53	6.166	2.6	1.234	9.982
Classes Per 1,000 School-Aged 1911	54	12.68	7.887	1.192	24.19
Classes Per 1,000 School-Aged 1921	54	11.70	3.701	4.956	20.45
Classes Per 1,000 School-Aged 1926	54	16.48	3.470	5.474	21.96
Teachers Per 1,000 School-Aged 1911	54	10.03	5.960	1.192	19.05
Teachers Per 1,000 School-Aged 1921	54	11.16	2.421	6.359	18.96
Teachers Per 1,000 School-Aged 1926	54	15.19	2.328	9.027	20.68

Table A.3: Primary enrollment at Prussian-Russian border 1911-1961

	(1)	(2)	(3)
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Prussian Side = 1	0.828***	0.829***	0.832***
	(0.019)	(0.018)	(0.018)
Observations	54	54	54
R-squared	0.995	0.996	0.996
Mean on Russian Side	0.164	0.164	0.164
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Prussian Side = 1	0.397***	0.402***	0.409***
	(0.052)	(0.036)	(0.038)
Observations	54	54	54
R-squared	0.784	0.905	0.913
Mean on Russian Side	0.562	0.562	0.562
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Prussian Side = 1	0.075*	0.080***	0.079***
	(0.038)	(0.029)	(0.029)
Observations	54	54	54
R-squared	0.098	0.651	0.652
Mean on Russian Side	0.711	0.711	0.711
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Prussian Side = 1	0.112***	0.111***	0.11***
	(0.026)	(0.025)	(0.026)
Observations	54	54	54
R-squared	0.36	0.512	0.515
Mean on Russian Side	0.961	0.961	0.961
Dep. Variable	Primary Enrollment 1961	Primary Enrollment 1961	Primary Enrollment 1961
Prussian Side = 1	-0.002	-0.001	-0.017
	(0.026)	(0.025)	(0.047)
Observations	53	53	53
R-squared	0.078	0.093	0.153
Mean on Russian Side	1.081	1.081	1.081
Distance, Distance*Prussian Side	Yes	Yes	Yes
Long./Lat., City Dummy	No	Yes	Yes
Geographic Controls	No	No	Yes

Notes: One-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.4: Primary enrollment at Austrian-Russian border 1911-1961

	(1)	(2)	(3)
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Austrian Side = 1	0.486***	0.498***	0.485***
	(0.055)	(0.055)	(0.055)
Observations	43	43	43
R-squared	0.902	0.921	0.926
Mean on Russian Side	0.194	0.194	0.194
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Austrian Side = 1	0.270***	0.275***	0.259***
	(0.053)	(0.051)	(0.046)
Observations	44	44	44
R-squared	0.716	0.827	0.856
Mean on Russian Side	0.554	0.554	0.554
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Austrian Side = 1	0.168***	0.169**	0.167**
	(0.057)	(0.068)	(0.072)
Observations	44	44	44
R-squared	0.304	0.354	0.386
Mean on Russian Side	0.627	0.627	0.627
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Austrian Side = 1	0.002	0.028	0.030
	(0.030)	(0.020)	(0.021)
Observations	44	44	44
R-squared	0.096	0.638	0.657
Mean on Russian Side	0.981	0.981	0.981
Dep. Variable	Primary Enrollment 1960/61	Primary Enrollment 1960/61	Primary Enrollment 1960/61
Austrian Side = 1	-0.007	-0.006	-0.001
	(0.015)	(0.016)	(0.016)
Observations	59	59	59
R-squared	0.070	0.141	0.189
Mean on Russian Side	1.062	1.062	1.062
Distance, Distance*Austrian Side	Yes	Yes	Yes
Long./Lat., City Dummy	No	Yes	Yes
Geographic Controls	No	No	Yes

Notes: One-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.5: Literacy at Prussian-Russian border 1921-1960

	(1)	(2)	(3)
Dep. Variable	Share of Literates 1921	Share of Literates 1921	Share of Literates 1921
Prussian Side = 1	0.255***	0.254***	0.248***
	(0.017)	(0.016)	(0.012)
Observations	54	54	54
R-squared	0.933	0.954	0.967
Mean on Russian Side	0.509	0.509	0.509
Dep. Variable	Share of Literates 1931	Share of Literates 1931	Share of Literates 1931
Prussian Side = 1	0.198***	0.196***	0.193***
	(0.017)	(0.012)	(0.012)
Observations	54	54	54
R-squared	0.935	0.974	0.977
Mean on Russian Side	0.517	0.517	0.517
Dep. Variable	Share of Literates 1960	Share of Literates 1960	Share of Literates 1960
Prussian Side = 1	0.025**	0.008	0.029***
	(0.011)	(0.014)	(0.011)
Observations	53	53	53
R-squared	0.474	0.655	0.884
Mean on Russian Side	0.936	0.936	0.936
Distance, Distance*Prussian Side	Yes	Yes	Yes
Long./Lat., City Dummy	No	Yes	Yes
Geographic Controls	No	No	Yes

Notes: One-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.6: Literacy at Austrian-Russian border 1921-1960

	(1)	(2)	(3)
Dep. Variable	Share of Literates 1921	Share of Literates 1921	Share of Literates 1921
Austrian Side = 1	0.270***	0.219***	0.200***
	(0.047)	(0.033)	(0.038)
Observations	44	44	44
R-squared	0.697	0.868	0.879
Mean on Russian Side	0.462	0.462	0.462
Dep. Variable	Share of Literates 1931	Share of Literates 1931	Share of Literates 1931
Austrian Side = 1	0.177***	0.132***	0.113***
	(0.037)	(0.022)	(0.024)
Observations	44	44	44
R-squared	0.597	0.830	0.851
Mean on Russian Side	0.491	0.491	0.491
Dep. Variable	Share of Literates 1960	Share of Literates 1960	Share of Literates 1960
Austrian Side = 1	0.057***	0.046***	0.045***
	(0.011)	(0.009)	(0.009)
Observations	59	59	59
R-squared	0.539	0.717	0.746
Mean on Russian Side	0.934	0.934	0.934
Distance, Distance*Austrian Side	Yes	Yes	Yes
Long./Lat., City Dummy	No	Yes	Yes
Geographic Controls	No	No	Yes

Notes: One-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.7: Enrollment at Prussian-Russian border 1911-1931, population controls

	(1)	(2)	(3)
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Prussian Side = 1	0.832***	0.846***	0.846***
	(0.018)	(0.028)	(0.03)
Share Jewish		0.430	0.430
		(0.296)	(0.3)
Share Protestant		0.059**	0.059
		(0.024)	(0.042)
Share Born Outside Partition 1921			-0.001
			(0.109)
Observations	54	54	54
R-squared	0.996	0.997	0.997
Dep.Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Prussian Side = 1	0.409***	0.357***	0.436***
	(0.038)	(0.067)	(0.080)
Share Jewish		-0.906	-0.837
		(0.767)	(0.794)
Share Protestant		-0.026	0.121
		(0.098)	(0.099)
Share Born Outside Partition 1921			-0.629***
			(0.212)
Observations	54	54	54
R-squared	0.913	0.917	0.924
Dep.Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Prussian Side = 1	0.079***	0.042	0.170***
	(0.029)	(0.044)	(0.052)
Share Jewish		-0.623	-0.479
		(0.461)	(0.469)
Share Protestant		0.054	0.316**
		(0.151)	(0.142)
Share Born Outside Partition 1921			-0.949***
			(0.185)
Observations	54	54	54
R-squared	0.652	0.665	0.759
Dep.Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Prussian Side = 1	0.110***	0.101***	0.079
	(0.026)	(0.034)	(0.050)
Share Jewish		-0.185	-0.217
		(0.343)	(0.353)
Share Protestant		-0.048	-0.099
		(0.183)	(0.208)
Share Born Outside Partition 1921			0.152
			(0.210)
Observations	54	54	54
R-squared	0.515	0.518	0.522
Geo, City Controls	Yes	Yes	Yes
Distance, Distance*Prussian Side, Long./Lat.	Yes	Yes	Yes

Notes: One-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.8: Enrollment at Austrian-Russian border 1911-1931, population controls

	(1)	(2)	(3)
Dep. Variable	Primary Enrollment 1911	Primary Enrollment 1911	Primary Enrollment 1911
Austrian Side = 1	0.481***	0.531***	0.533***
	(0.057)	(0.053)	(0.053)
Share Jewish		0.561*	0.557*
		(0.306)	(0.325)
Share Protestant		5.874**	0.566**
		(2.157)	(2.129)
Share Born Outside Partition 1921			0.473
			(0.833)
Observations	44	44	44
R-squared	0.913	0.937	0.938
Dep. Variable	Primary Enrollment 1921	Primary Enrollment 1921	Primary Enrollment 1921
Austrian Side = 1	0.259***	0.165***	0.165***
	(0.046)	(0.049)	(0.043)
Share Jewish		-1.163***	-1.300***
		(0.299)	(0.255)
Share Protestant		-0.303	-0.529
		(4.214)	(4.176)
Share Born Outside Partition 1921			1.546***
			(0.507)
Observations	44	44	44
R-squared	0.856	0.890	0.902
Dep. Variable	Primary Enrollment 1926	Primary Enrollment 1926	Primary Enrollment 1926
Austrian Side = 1	0.167**	0.136*	0.136*
	(0.072)	(0.077)	(0.078)
Share Jewish		-0.413	-0.431
		(0.667)	(0.688)
Share Protestant		0.164	-0.051
		(3.789)	(3.894)
Share Born Outside Partition 1921			0.246
			(1.043)
Observations	44	44	44
R-squared	0.386	0.393	0.394
Dep. Variable	Primary Enrollment 1932	Primary Enrollment 1932	Primary Enrollment 1932
Austrian Side = 1	0.030	0.034	0.036
	(0.021)	(0.029)	(0.026)
Share Jewish		0.055	0.024
		(0.244)	(0.226)
Share Protestant		0.676	-0.321
		(2.019)	(2.371)
Share Born Outside Partition 1921			0.656*
			(0.336)
Observations	44	44	44
R-squared	0.657	0.659	0.683
Geo, City Controls	Yes	Yes	Yes
Distance, Distance*Austrian Side, Long./Lat.	Yes	Yes	Yes

Notes: One-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.9: Literacy at Prussian-Russian border 1921-1931, population controls

	(1)	(2)	(3)
Dep. Variable	Share of Literates 1921	Share of Literates 1921	Share of Literates 1921
Prussian Side = 1	0.248***	0.263***	0.264***
	(0.012)	(0.020)	(0.024)
Share Jewish		0.219	0.221
		(0.253)	(0.257)
Share Protestant		-0.011	-0.008
		(0.038)	(0.049)
Share Born Outside Partition 1921			-0.014
			(0.109)
Observations	54	54	54
R-squared	0.967	0.968	0.968
Mean on Russian Side	0.509	0.509	0.509
Dep. Variable	Share of Literates 1931	Share of Literates 1931	Share of Literates 1931
Prussian Side = 1	0.193***	0.212***	0.240***
	(0.012)	(0.015)	(0.018)
Share Jewish		0.385**	0.426**
		(0.180)	(0.188)
Share Protestant		0.014	0.078
		(0.081)	(0.061)
Share Born Outside Partition 1921			-0.191***
			(0.061)
Observations	54	54	54
R-squared	0.977	0.979	0.982
Mean on Russian Side	0.517	0.517	0.517

Notes: One-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A.10: Literacy at Austrian-Russian border 1921-1931, population controls

	(1)	(2)	(3)
Dep. Variable	Share of Literates 1921	Share of Literates 1921	Share of Literates 1921
Austrian Side = 1	0.200***	0.194***	0.194***
	(0.038)	(0.055)	(0.052)
Share Jewish		-0.043	-0.125
		(0.321)	(0.330)
Share Protestant		2.256	2.121
		(1.535)	(1.549)
Share Born Outside Partition 1921			0.927
			(0.569)
Observations	44	44	44
R-squared	0.879	0.882	0.890
Mean on Russian Side	0.462	0.462	0.462
Dep. Variable	Share of Literates 1931	Share of Literates 1931	Share of Literates 1931
Austrian Side = 1	0.113***	0.134***	0.137***
	(0.024)	(0.035)	(0.032)
Share Jewish		0.294	0.255
		(0.249)	(0.237)
Share Protestant		2.586	1.340
		(1.855)	(2.335)
Share Born Outside Partition 1921			0.821*
			(0.406)
Observations	44	44	44
R-squared	0.851	0.865	0.879
Mean on Russian Side	0.491	0.491	0.491

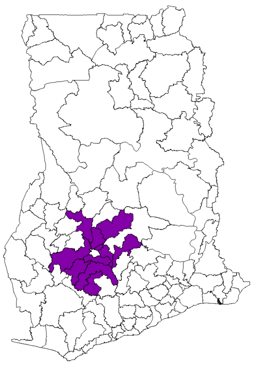
Notes: One-dimensional RDD. Bandwidth 65 km. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Appendix B

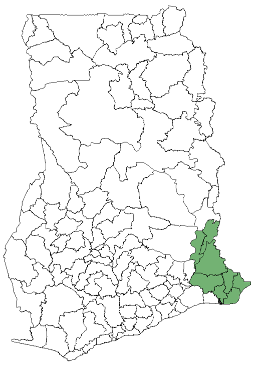
Appendix to Chapter 2

B.1 Figures

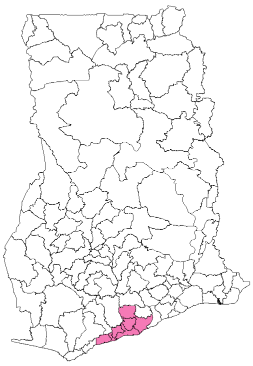
Figure B.1: Ethnic groups of Ghana



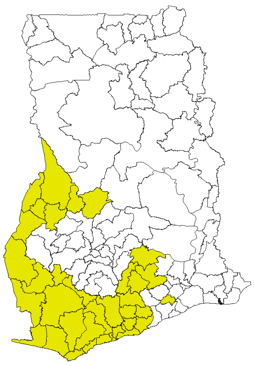
(a) Asante



(b) Ewe

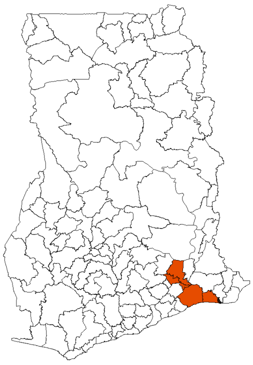


(c) Fante

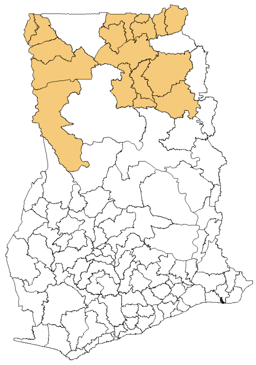


(d) Other Akan groups

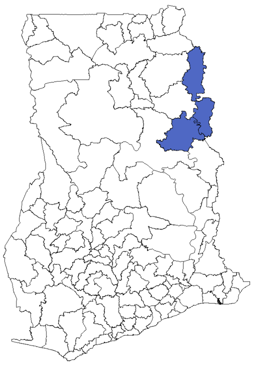
Figure B.1: Ethnic groups of Ghana (continued)



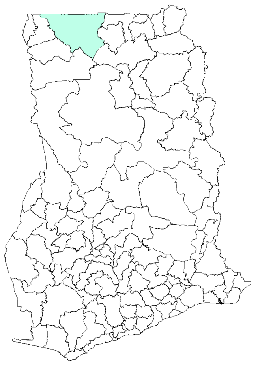
(e) Ga-Adangbe



(f) Mole-Dagbani



(g) Gurma



(h) Grusi

B.2 Tables

Table B.1: Summary statistics for the time period 1992-2008

Variable	(1) N	(2) Mean	(3) SD	(4) Min	(5) Max
Log Nighttime Lights	1,760	-1.156	1.941	-4.605	4.067
Share of Ewe	1,760	11.92	21.64	0	91.47
Share of Asante	1,760	11.74	20.68	0.0511	88.24
Share of Akan (w/o Asante)	1,760	31.72	28.99	0.386	90.52
Share of Fante	1,760	8.933	19.50	0	84.54
Share of Ga-Adangbe	1,760	5.979	14.56	0.0256	81.79
Share of Mole-Dagbani	1,760	17.37	26.41	0.197	92.92
Share of Guan	1,760	4.392	8.883	0.0964	48.77
Share of Gurma	1,760	4.444	11.73	0.0685	65.19
Share of Grusi	1,760	2.928	8.029	0.0340	75.96
Share of Other Groups	1,760	1.508	3.003	0.0132	17.08
Share of Undefined Groups	1,760	7.260	4.324	0.794	28.29
Vote Margin	1,760	36.76	25.54	0.119	98.23

Appendix C

Appendix to Chapter 3

C.1 Tables

Table C.1: Summary statistics for the sample at the district level

Country	Variable	N	Mean	SD	Min	Max
Bangladesh	Electrification	128	0.384	0.203	0.0749	0.971
	Years of Schooling	128	3.724	0.758	2.278	6.245
	Log Nighttime Lights	128	0.551	0.942	-1.880	3.084
	Log Population Per Pixel	128	6.572	0.638	3.928	8.743
Cambodia	Electrification	316	0.186	0.244	0	0.996
	Years of Schooling	316	2.971	1.214	0.331	7.321
	Log Nighttime Lights	316	-3.207	2.250	-4.605	4.119
	Log Population Per Pixel	316	4.599	1.589	0.300	10.78
Ecuador	Electrification	154	0.883	0.0971	0.510	0.995
	Years of Schooling	154	6.070	1.256	3.133	9.663
	Log Nighttime Lights	154	0.948	1.286	-2.848	3.694
	Log Population Per Pixel	154	3.919	1.122	0.584	7.262
Ghana	Electrification	204	0.404	0.221	0.0404	0.950
	Years of Schooling	204	4.003	1.536	0.789	8.024
	Log Nighttime Lights	204	-0.619	1.866	-4.605	4.115
	Log Population Per Pixel	204	4.568	1.120	2.239	8.967
Malawi	Electrification	446	0.0806	0.151	0	0.921
	Years of Schooling	446	3.947	1.525	1.271	8.828
	Log Nighttime Lights	446	-1.564	2.929	-4.605	4.086
	Log Population Per Pixel	446	5.289	1.436	2.163	9.238
Mali	Electrification	94	0.0720	0.102	0.000388	0.757
	Years of Schooling	94	0.825	0.597	0.162	4.282
	Log Nighttime Lights	94	-3.681	1.400	-4.605	3.814
	Log Population Per Pixel	94	2.597	1.542	-1.761	8.732
Zambia	Electrification	140	0.148	0.176	0	0.641
	Years of Schooling	140	4.164	1.245	1.831	7.711
	Log Nighttime Lights	140	-2.445	1.941	-4.605	3.740
	Log Population Per Pixel	140	2.674	1.314	0.578	8.050

Table C.2: Summary statistics for the sample at the regional level

Country	Variable	N	Mean	SD	Min	Max
Bangladesh	Years of Schooling	12	3.871	0.563	2.804	4.620
	Electrification	12	0.414	0.159	0.197	0.675
	Log Nighttime Lights	12	0.663	0.481	-0.229	1.456
	Log Population Per Pixel	12	6.541	0.243	6.216	7.081
Cambodia	Years of Schooling	46	2.898	1.024	0.829	6.337
	Electrification	46	0.205	0.191	0.0310	0.948
	Log Nighttime Lights	46	-2.987	1.739	-4.605	3.177
	Log Population Per Pixel	46	3.997	1.513	0.651	7.998
Ecuador	Years of Schooling	28	6.521	0.919	4.887	7.935
	Electrification	28	0.873	0.104	0.612	0.977
	Log Nighttime Lights	28	0.358	1.304	-2.702	2.075
	Log Population Per Pixel	28	3.340	1.276	0.584	4.851
Ghana	Years of Schooling	20	4.145	1.629	1.616	7.462
	Electrification	20	0.457	0.213	0.123	0.879
	Log Nighttime Lights	20	-0.476	1.490	-3.011	2.716
	Log Population Per Pixel	20	4.498	0.972	3.089	6.830
Malawi	Years of Schooling	48	3.650	0.927	2.080	5.982
	Electrification	48	0.0381	0.0481	0.00612	0.289
	Log Nighttime Lights	48	-1.767	1.049	-3.712	1.690
	Log Population Per Pixel	48	4.556	0.696	3.122	6.029
Mali	Years of Schooling	16	1.200	1.050	0.395	4.282
	Electrification	16	0.135	0.190	0.0165	0.757
	Log Nighttime Lights	16	-2.634	2.504	-4.549	3.814
	Log Population Per Pixel	16	2.917	2.515	-0.260	8.732
Zambia	Years of Schooling	16	4.549	1.200	3.108	6.972
	Electrification	16	0.186	0.176	0.0304	0.567
	Log Nighttime Lights	16	-2.113	1.503	-3.850	0.555
	Log Population Per Pixel	16	2.627	0.960	1.131	4.383

Bibliography

Acemoglu, D., Johnson, S. and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, **91** (5), 1369–1401.

Afrobarometer (2015). Afrobarometer data, Ghana, round 6. Available at http://www.afrobarometer.org.

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, **24** (2), 3–30.

Austin, G. (2008). The reversal of fortune thesis and the compression of history: Perspectives from African and comparative economic history. *Journal of International Development*, **20** (8), 996–1027.

Banful, A. B. (2011). Do formula-based intergovernmental transfer mechanisms eliminate politically motivated targeting? Evidence from Ghana. *Journal of Development Economics*, **96** (2), 380–390.

Becker, S. O., Boeckh, K., Hainz, C. and Woessmann, L. (2016). The empire is dead, long live the empire! Long-run persistence of trust and corruption in the bureaucracy. *The Economic Journal*, **126** (590), 40–74.

—, Cinnirella, F., Hornung, E. and Woessmann, L. (2014). iPEHD - the ifo Prussian economic history database. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, **47** (2), 57–66.

—, Hornung, E. and Woessmann, L. (2011). Education and catch-up in the industrial revolution. *American Economic Journal: Macroeconomics*, **3** (3), 92–126.

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

Bhandari, L. and Roychowdhury, K. (2011). Night lights and economic activity in India: A study using DMSP-OLS night time images. *Proceedings of the Asia-Pacific advanced network*, **32**, 218–236.

Bickenbach, F., Bode, E., Nunnenkamp, P. and Söder, M. (2016). Night lights and regional GDP. *Review of World Economics*, **152** (2), 425–447.

Blanke, R. (1993). *Orphans of Versailles*. University Press of Kentucky.

Bukowski, P. (2016). How history matters for student performance. Lessons from the Partitions of Poland. Working paper.

Bureau der K.K. Statistischen Zentralkommission (1914). Die Bevölkerung nach der Gebürtigkeit, Religion und Umgangssprache in Verbindung mit dem Geschlechte, nach dem Bildungsgrade und Familienstande; Die körperlichen Gebrechen; Die soziale Gliederung der Haushaltungen. *Österreichische Statistik*, **1** (2), 1–103.

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A. and Padró i Miquel, G. (2015). The value of democracy: Evidence from road building in Kenya. *The American Economic Review*, **105** (6), 1817–1851.

Caicedo, F. V. (2015). The mission: Human capital transmission, economic persistence and culture in South America. Working paper.

Cantoni, D. and Yuchtman, N. (2013). The political economy of educational content and development: Lessons from history. *Journal of Development Economics*, **104**, 233–244.

Casey, K. (2015). Crossing party lines: The effects of information on redistributive politics. *American Economic Review*, **105** (8), 2410–48.

Chen, X. (2015). Explaining subnational infant mortality and poverty rates: What can we learn from night-time lights? *Spatial Demography*, **3** (1), 27–53.

— and Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, **108** (21), 8589–8594.

Cinnirella, F. and Schueler, R. M. (2017). The cost of decentralization: Linguistic polarization and the provision of education. Mimeo.

Comin, D., Easterly, W. and Gong, E. (2010). Was the wealth of nations determined in 1000 BC? *American Economic Journal: Macroeconomics*, **2** (3), 65–97.

Cox, G. W. and McCubbins, M. D. (1986). Electoral politics as a redistributive game. *The Journal of Politics*, **48** (2), 370–389.

Davies, N. (2005). *God's Playground: A History of Poland: Volume II: 1795 to the Present*, vol. 2. Oxford University Press.

De Luca, G., Hodler, R., Raschky, P. A. and Valsecchi, M. (2016). Ethnic favoritism: An axiom of politics? Working paper.

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

Dittmar, J. E. and Meisenzahl, R. R. (2016). State capacity and public goods: Institutional change, human capital, and growth in early modern Germany. Working paper.

Dixit, A. and Londregan, J. (1996). The determinants of success of special interests in redistributive politics. *The Journal of Politics*, **58** (4), 1132–1155.

Dobosiewicz, S. (1970). Reforma szkolna 1961 roku: Geneza i założenia. *Rozprawy z Dziejów Oświaty*, **13**, 107–162.

Dupraz, Y. (2017). French and British colonial legacies in education: Evidence from the partition of Cameroon. University of Warwick. Working Paper Series. No. 333.

Eberhardt, P. (2011). *Political migrations on Polish territories (1939-1950)*. Instytut Geografii i Przestrzennego Zagospodarowania im. Stanisława Leszczyckiego PAN.

Elbers, C., Lanjouw, J. O. and Lanjouw, P. (2003). Micro–level estimation of poverty and inequality. *Econometrica*, **71** (1), 355–364.

Elvidge, C. D., Baugh, K. E., Zhizhin, M. and Hsu, F.-C. (2013). Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proceedings of the Asia-Pacific Advanced Network*, **35**, 62–69.

Falski, M. (1925). Materiały do projektu sieci szkół powszechnych na terenie województw warszawskiego, łódzkiego, kieleckiego, białostockiego, m. st. Warszawy.

Franck, R. and Rainer, I. (2012). Does the leader's ethnicity matter? Ethnic favoritism, education, and health in sub-Saharan Africa. *American Political Science Review*, **106** (2), 294–325.

Galor, O. and Özak, Ö. (2016). The agricultural origins of time preference. *The American Economic Review*, **106** (10), 3064–3103.

Gawryszewski, A. (2005). *Ludność Polski w XX wieku*. Instytut Geografii i Przestrzennego Zagospodarowania im. Stanisława Leszczyckiego PAN.

Gelman, A. and Imbens, G. (2016). Why high-order polynomials should not be used in regression discontinuity designs. Working paper.

Gennaioli, N., La Porta, R., Lopez-de Silanes, F. and Shleifer, A. (2013). Human capital and regional development. *The Quarterly Journal of Economics*, **128** (1), 105–164.

Glaeser, E. L., La Porta, R., Lopez-de Silanes, F. and Shleifer, A. (2004). Do institutions cause growth? *Journal of Economic Growth*, **9** (3), 271–303.

Grosfeld, I., Rodnyansky, A. and Zhuravskaya, E. (2013). Persistent antimarket culture: A legacy of the pale of settlement after the Holocaust. *American Economic Journal: Economic Policy*, **5** (3), 189–226.

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

Grossman, H. (1925). Struktura spoleczna i gospodarcza Ks. Warszawskiego na podstawie spisow 1808-1810 roku. *Kwartalnik Statystyczny*, **2**, 1–224.

GUS (1922). *Revue mensuelle de statistique*, vol. 5.

— (1927). *Statystyka Polski: Pierwszy Powszechny Spis Rzeczypospolitej Polski.*

— (1934). *Statystyka szkolnictwa 1932/33.*

— (1938). *Statystyka Polski: Drugi Powszechny Spis.*

— (1955). *Narodowy spis powszechny z dnia 3 grundnia 1950 r.*

— (1962). *Statystyka szkolnictwa.*

— (1965). *Spis powszechny z dnia 6 grundnia 1960 r.*

Hanushek, E. A. and Woessmann, L. (2012a). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, **17** (4), 267–321.

— and — (2012b). Schooling, educational achievement, and the Latin American growth puzzle. *Journal of Development Economics*, **99** (2), 497–512.

Hargrove, J. L. (2015). *The Political Economy of the Interior Gold Coast: The Asante and the Era of Legitimate Trading, 1807–1875*. Lexington Books.

Henderson, J. V., Storeygard, A. and Weil, D. N. (2012). Measuring economic growth from outer space. *The American Economic Review*, **102** (2), 994–1028.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25** (15), 1965–1978.

Hodler, R. and Raschky, P. A. (2014). Regional favoritism. *The Quarterly Journal of Economics*, **129** (2), 995–1033.

Hornung, E. (2014). Immigration and the diffusion of technology: The Huguenot diaspora in Prussia. *The American Economic Review*, **104** (1), 84–122.

Huang, Q., Yang, X., Gao, B., Yang, Y. and Zhao, Y. (2014). Application of DMSP/OLS nighttime light images: A meta-analysis and a systematic literature review. *Remote Sensing*, **6** (8), 6844–6866.

Huntington, S. P. (1991). Democracy's third wave. *Journal of Democracy*, **2** (2), 12–34.

Ichino, N. and Nathan, N. L. (2013). Crossing the line: Local ethnic geography and voting in Ghana. *American Political Science Review*, **107** (2), 344–361.

Jedwab, R. and Moradi, A. (2016). The permanent effects of transportation revolutions in poor countries: Evidence from Africa. *The Review of Economics and Statistics*, **98** (2), 268–284.

Kancelaria Sejmu Rzeczypospolitej Polskiej (2015). Internetowy system aktow prawnych.

Königlich Preussisches Statistisches Landesamt in Berlin (1912). Das niedere Schulwesen in Preußen, 1911. II. Teil. Die öffentlichen Volksschulen nach Kreisen. **231**.

Kramon, E. and Posner, D. N. (2016). Ethnic favoritism in education in Kenya. *Quarterly Journal of Political Science*, **11** (1), 1–58.

Krzesniak-Firlej, D., Majewski, S. and Firlej, W. (2014). *Organizacja szkolnictwa w II Rzeczypospolitej 1918-1939*. Uniwersytet Jana Kochanowskiego w Kielcach.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, **48** (2), 281–355.

Lindbeck, A. and Weibull, J. W. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice*, **52** (3), 273–297.

Lindberg, S. I. (2013). Have the cake and eat it: The rational voter in Africa. *Party Politics*, **19** (6), 945–961.

— and Morrison, M. K. (2005). Exploring voter alignments in Africa: Core and swing voters in Ghana. *The Journal of Modern African Studies*, **43** (4), 565–586.

Markevich, A. and Zhuravskaya, E. (2017). The economic effects of the abolition of serfdom: Evidence from the Russian empire. Forthcoming *The American Economic Review*.

Michalopoulos, S. and Papaioannou, E. (2013a). National institutions and subnational development in Africa. *The Quarterly Journal of Economics*, **129** (1), 151–213.

— and — (2013b). Pre-colonial ethnic institutions and contemporary African development. *Econometrica*, **81** (1), 113–152.

Minister Wyznan Religijnych i Oswiecenia Publicznego (1927). Szkoły powszechne rzeczypospolitej polskiej w roku szkolnym 1925/26–stan z dnia 1 grudnia 1925 roku.

Ministerstwo Oświaty (1946). Szkolnictwo w liczbach 1944/45.

Minnesota Population Center (2017). Integrated public use microdata series, international: Version 6.5 [dataset].

MPIDR and CGG (2011). Mpidr population history GIS collection (partly based on Hubatsch and Klein 1975 ff.).

MPIDR and CGG (2012a). Mpidr population history GIS.

MPIDR and CGG (2012b). Mpidr population history GIS collection (slightly modified version of a gis-file by Rumpler and Seger 2010).

Nordhaus, W. and Chen, X. (2015). A sharper image? Estimates of the precision of nighttime lights as a proxy for economic statistics. *Journal of Economic Geography*, **15** (1), 217–246.

Nunn, N. (2009). The importance of history for economic development. *Annual Review of Economics*, **1** (1), 65–92.

— and Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *The American Economic Review*, **101** (7), 3221–3252.

Persson, T. and Tabellini, G. (2009). Democratic capital: The nexus of political and economic change. *American Economic Journal: Macroeconomics*, **1** (2), 88–126.

Piketty, T. (2003). Income inequality in France, 1901–1998. *Journal of Political Economy*, **111** (5), 1004–1042.

Pilata, T. (ed.) (1913). *Podręcznik statystyki galicyi*, vol. IX. Krajowe biuro statystyczne.

Pischke, S. (2007). Lecture notes on measurement error.

Pokrowskoho, B. I. (ed.) (1914). *One-day census of primary schools in the Empire, performed on January 18, 1911 (translated)*, vol. X: Warsaw academic district. Ministry of National Education.

QGIS Development Team (2017). *QGIS Geographic Information System.* Open Source Geospatial Foundation.

Ramankutty, N. and Foley, J. A. (1999). Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global biogeochemical cycles*, **13** (4), 997–1027.

Rigobon, R. and Stoker, T. M. (2007). Estimation with censored regressors: Basic issues. *International Economic Review*, **48** (4), 1441–1467.

— and — (2009). Bias from censored regressors. *Journal of Business & Economic Statistics*, **27** (3), 340–353.

Squicciarini, M. P. and Voigtländer, N. (2015). Human capital and industrialization: Evidence from the age of enlightenment. *The Quarterly Journal of Economics*, **130** (4), 1825–1883.

Tarozzi, A. and Deaton, A. (2009). Using census and survey data to estimate poverty and inequality for small areas. *The Review of Economics and Statistics*, **91** (4), 773–792.

Trenkler, C. and Wolf, N. (2005). Economic integration across borders: The Polish interwar economy 1921–1937. *European Review of Economic History*, **9** (2), 199–231.

Ward, W. E. F. (1966). *A History of Ghana*. 3rd edn.

Wittenberg, J. (2015). Conceptualizing historical legacies. *East European Politics and Societies*, **29** (2), 366–378.

Wolf, N. (2005). Path dependent border effects: The case of Poland's reunification (1918–1939). *Explorations in Economic History*, **42** (3), 414–438.

Wysokinska, A. A. (2017). Institutions or culture? Lessons for development from two natural experiments of history. Working paper.

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

Datum: München, 04.05.2018

Unterschrift: Backhaus, Andreas