

The Effect of Napoleon Bonaparte on Education Systems and its Consequences: The Education Revolution on Horseback¹

Abstract: Sociological literature argues that the threat for war with and invasion by the French around the 1800s induced European countries to introduce mass public education systems. On the basis of this theory, the aim of this chapter is twofold: First, it empirically establishes whether the proposed mechanism had an influence on the levels of tracking and standardization in European education systems. Second, it studies whether tracking has an effect on student performance by instrumenting tracking by the political pressure caused by the Napoleonic wars. Especially for tracking, a relation between education systems and the pressure from Napoleon is found. For standardization, no or a much weaker relation is found. Using Napoleon as an instrument for tracking leads to a strong first stage. The resulting analyses show a consistent positive effect between the level of tracking and student performance.

¹. I would like to thank Jaap Dronkers and Sjoerd Karsten for bringing the paper of Ramirez and Boli (1987) under my attention, and the Forschungsdatenzentrum am Institut zur Qualitätsentwicklung im Bildungswesen for providing me with the extended PISA data for Germany. I further would like to thank Francisco Ramirez, David Figlio, Bas ter Weel, and participants of the International Workshop on Applied Economics of Education 2013, the Workshop on Comparative systems of educational and political systems at Stanford University, and ESPE 2014 for valuable comments.

“Our view is that European states became engaged in authorizing, funding and managing mass schooling as a part of the endeavor to construct a unified national policy. [...] External challenges [...] were important stimuli to state action in education [...]” (Ramirez and Boli, 1987, p. 3)

“We show that state educational systems did not originate in the most dominant countries, contrary to the widespread assumption that educational innovations inevitably start there [...] Rather, the most dominant powers were able to resist the system wide pressures favoring mass education [...].” (Ramirez and Boli, 1987, p. 4)

2.1 Introduction

Looking at the education systems in Europe (see Figure 1 in Section 2.2), patterns can be distinguished. For instance, it is easily seen that the Scandinavian countries and the UK have very similar comprehensive education systems with only one track available to students and late selection. Contrary to that, Germany, the Netherlands, Belgium, Switzerland, and Austria have highly differentiated systems with a high number of tracks and, in Germany and Austria, selection as early as 10 years old. The geographical spread of standardization of inputs (administrative standardization and content related standardization) and outputs (central exit exams) is less clear. Ramirez and Boli (1987) propose a theory that points to an explanation for the geographical spread of differentiation found in European education systems: the threat of war with and annexation by the French in the 1800s induced European countries to create mass public education systems. This chapter investigates whether the proposed mechanism had an effect on levels of tracking and standardization in European education systems and what the implementations of this are.

The aim of this chapter is twofold: First, it takes the theory by Ramirez and Boli (1987) and empirically tests whether the proposed mechanism also had effect on levels of tracking and standardization in European education systems. In Section 2.2 it will be shown that the pressure from France influenced European educational systems and that this influence can still be seen today, although not for all aspects of education systems. The political pressure from France, as theorized of Ramirez and Boli (1987), is related to the level of differentiation, both seen by the number of tracks available to fifteen-year-old students and by the age of first selection. The mechanism seems to have had less influence on content related standardization, and none on central exams and administrative standardization.

The second aim of this chapter, discussed in Section 2.3, is to explain current student performance by building on the theory by Ramirez and Boli (1987) and using the historical variation in education systems. The pressure from France is used as an instrument to explain education system characteristics to alleviate the endogeneity

between education systems and student performance. Naturally, the French influence in the 1800s influenced many aspects of European societies and thus these analyses should not be regarded as pure causal analyses, but as an alternative to existing studies to describe education systems and its consequences. Section 2.5 provides more information on possible other influences of Napoleon and other threats to the identification. Given the considerations there, tracking has a strong first stage and positive effects are found for the number of tracks in a country. Negative, but insignificant, effects are found for the age of first selection.

In the 18th and 19th century most European countries passed their first widespread compulsory schooling laws and centralized education systems were set up. Before this period, schooling was not available to all and the educational institutions were in non-state control, controlled especially by the church(es) and local authorities. The Enlightenment in the 18th century paved the way for mass public education, since schooling would ensure advancement of knowledge so sought after by this movement. However, it was other, more political, events which induced the actual introduction of the educational systems in Europe. The late 18th century started, partly as a consequence of the political success of the Enlightenment in France, a volatile period in European history, best seen by the French Revolution of 1789 and its impact on the continent. After Napoleon claimed power in France in 1799, the Napoleonic wars dominated the European continent for almost two decades. The French army conquered large areas of Europe until Napoleon was weakened by the failed Russian invasion in 1812 and finally defeated in 1815 in the battle of Waterloo. In the period in between, almost all European powers were invaded or threatened to be invaded by Napoleon.

As laid out by Ramirez and Boli (1987), the threat of war with and annexation by the French induced European countries to create a national identity among their population. An important part of this nation building was the introduction of mass public educational systems by the national state, intended mainly to educate the future generations in the cultural and political traditions of the nation. Creating a national identity in war times is important for states since it increases the public's (physical and mental) willingness to pay and fight for the survival of their own nation in order to remain independent.

This chapter is structured as follows: The next section gives an introduction to the history of French influence in the 1800s and elaborates on the theory by Ramirez and Boli (1987). Section 2.3 presents evidence on the influence of Napoleon on a number of education system characteristics. Section 2.4 uses the proposed theory by Ramirez and Boli (1987) to employ an instrumental variable method to alleviate the endogeneity between education system characteristics and student performance. Section 2.5 gives attention to the exclusion restriction. Section 2.6 summarizes and concludes.

2.2 Napoleon and education systems

This section first gives a short historical overview of Napoleon and the Napoleonic Wars. Then it provides a discussion of the theory proposed by Ramirez and Boli (1987). This theory relates external pressure to the introduction of mass public education systems. The different stages of this process are discussed, and examples are given. Finally, using data on education systems, the relation is shown between political pressure from Napoleon and education systems.

2.2.1 Political pressure from the Napoleonic wars²

The external challenges, referred to by Ramirez and Boli (1987), were in most cases (political) pressure arising from the French army, and later and foremost from Napoleon. After the turmoil of the French Revolution and its aftermath, Napoleon Bonaparte, a general who just secured victories for France in Italy, Austria, and Egypt³, was called in to support a coup d'état by the National Convention, consisting of the bourgeois who wanted to overthrow the everlasting power of the clergy and the nobility. However, Napoleon took power for himself in November 1799 and declared himself First Consul in 1800.

Almost as soon as Napoleon took power, he started invading neighboring countries and annexed Piedmont in Italy, the Rhineland in Germany and Belgium, appointed himself president of the Republic of Italy and Mediator of the Swiss Confederation. In 1802 the peace treaty of Amiens was signed by France, England, Holland and Spain stopping combat and reaffirming the dependence of the earlier conquered states to France. But the peace did not last long as Napoleon kept intervening in neighboring countries and setting up strategic positions against England, the most dominant (naval) power at the time. Soon after his coronation to Emperor of France in 1804, Napoleon expanded his scope and also annexed the Ligurian Republic (1805, Italy, south of Piedmont). In response, in 1805 the Russian Empire and England formed an alliance against France and its expansionary urges, which Austria (the Habsburger Empire), Sweden, and Naples later also joined (see Figure 2 for a division of Europe in the different empires). Napoleon formed alliances with Spain and Bavaria in return. Despite the alliance with Russia and England, Austria was conquered and was given harsh settlement terms and had to give up a number of territories. France became the dominant power on the continent.⁴

2. This paragraph is based on Santon and MacKay (2010) and the first and Chapter 13 of Grab (2003), *The formation of the Napoleonic Empire and The collapse of the Napoleonic Empire*.

3. Napoleons journey to Egypt did not end as well as it started. After Napoleon conquered Egypt, and with it was able to frustrate the trade routes of England, he lost most of its fleet in the Battle of Aboukir and was unable to return home. Napoleon ruled Egypt for over about a year before leaving behind his army and returning to France in 1799.

4. Although Napoleon had less luck overseas, which can be illustrated by his defeat over England at the (sea) Battle of Trafalgar in 1805.

France was also able to force Prussia into submission: Prussia had to end trade with England but received the conquered region of Hanover from France in return. However, Prussia sided with Russia and sought combat with France but was defeated in 1806. Napoleon conquered Berlin and later also invaded Prussia's Poland and established the Duchy of Warsaw. On his way east, Napoleon met the combined forces from Prussia and Russia in 1807 but overcame. Russia had to acknowledge Napoleon's brothers as rulers of Naples, Holland, and Westphalia and in return Napoleon would leave the Baltics states and Turkey alone. Russia took its chance and occupied Finland from Sweden in 1808. Prussia had to give up territories to the Kingdom of Westphalia and the Duchy of Warsaw, which were under Napoleons influence, pay damages, and accept limitations on its army. Now also Germany and the east were under Napoleons control and Russia had given up its resistance to the French.

In 1808 while the French army was mostly tied up in Spain, Austria started to make preparations to go to war against France. The war was fought in southern Germany and Austria and in Italy and Poland. Both sides won and lost battles, and even though England came to Austria's defense in the Netherlands, in 1809 Austria had to concede to France once again. France received the Illyrian Provinces ("Croatia"), while also Bavaria and the Duchy of Warsaw received lands from Austria. Austria had to pay damages and had to obey limitations on its army. Later Napoleon married the daughter of Francis I, the emperor of Austria, and thus formed an alliance between both houses. Pope Pius VII opposed the marriage, as he had not approved Napoleon's divorce from his former wife Josephine, and also refused to prevent England ships from using its ports, and thus Napoleon marched south to Rome. In 1809 Napoleon had taken over control of Rome and the remaining Papal States and exiled Pope Pius.

Under Napoleons reign the French Empire grew substantially: At the height of the French Empire in 1810 it covered an area of 750,000 square kilometers (293,000 square miles) as compared to 543,965 square kilometer (210,026 square miles) for Metropolitan France today. But not all territories were under French control. There were "allied countries whose territory was expanded by Napoleon in some cases, and continued to be governed by their native rulers", like Bavaria and Sweden. Grab (2003) calls these the *pays allies*. The conquered or annex countries were either "countries annexed to France and directly ruled by Napoleon" (*pays réunis*), like Belgium and Piedmont, or "satellite states that were entrusted to French rulers" (*pays conquis*), like Spain and Switzerland. Figure 3 shows how Europe was divided under Napoleons reign.

Napoleon became less popular in France over the year with the continued fighting, the exile of Pope Pius VII in 1809, and when due to the blockade of England and other countries opposed to Napoleon a shortage of materials arose. Also the alliance with Russia proved to be weak. When Russia refused to keep up the blockade towards England, both Russia and France made preparations for war. Alliances were sought and soldiers were readied. Napoleon advanced into Russia but lost many men along the way due to heat, hunger, and disease. Napoleon took Moscow in 1812 but in doing so the supply lines to the front line became too long and supplies grew thin while the cold was setting in. Napoleon retreated but lost even more men along the way back, this time due to cold and Russian troops.

Napoleon returned to Paris to raise another army. This time however, Russia, Prussia, England, Sweden and Austria united against France, and also France's long-time ally Bavaria reluctantly joined this coalition later on. In Spain and Italy, Napoleon also faced problems and he finally abdicated when Paris was taken in April 1814. Napoleon was sent to Elba but he came back to Paris in 1815 for the Hundred Days until he was finally defeated in the battle of Waterloo. Napoleon was exiled to St. Helena and died there in 1821.

2.2.2 *A theory on the origins of education systems*⁵

In their 1987 paper *The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization* Ramirez and Boli present their theory on how external pressure led to similar events in multiple countries which eventually led to the introduction of mass public education systems in the nineteenth century. Their theory can best be described by the two quotes at the beginning of the chapter of which the first is: "Our view is that European states became engaged in authorizing, funding, and managing mass schooling as a part of the endeavor to construct a unified national policy. [...] External challenges [...] were important stimuli to state action in education [...]" (Ramirez and Boli, 1987, p. 3). Napoleon can be seen as one of the most pronounced external challenges of this era.

Ramirez and Boli (1987) illustrate their theory by describing this process in seven European regions: Prussia, Denmark, Austria, Sweden, Italy, France, and England. A summary of these case studies is provided in Appendix A. Sweden, France, and England experienced a general call from within for mandatory, universal and free education later than other countries. Prussia and Denmark already made early moves towards education for all. Italy is an interesting case since the South acted like France and England, while the North made, like Prussia and Denmark, earlier steps towards a mass public education system. This supports the theory by Ramirez and Boli (1987)

5. This section is based upon Ramirez and Boli (1987).

that regions that were under external pressure (by Napoleon in most cases) were induced to introduce mass public education systems while “the most dominant powers were able to resist the system wide pressures favoring mass education [...]” (Ramirez and Boli, 1987, p. 4).

Although national events and characteristics naturally played a large role in the origins of education systems, Ramirez and Boli (1987) focus only on “transnational similarities in the institutional character of state educational systems” (p. 2). Each of the regions discussed by Ramirez and Boli showed the same pattern in response to external challenges. These patterns follow 4 steps (p. 9). First there was a “declaration of a national interest in mass education”. In Germany this can be seen by an address by the German philosopher Johann Gottlieb Fichte in 1807 which stated that “universal state-directed, compulsory education would teach all Germans to be good Germans and would prepare them to play whatever role [...] fell to them in helping the state reassert Prussian power” (p. 5). In Denmark this can be seen by the claims between 1807 and 1814 by N.F.S. Grundvig, a clergyman, that education “[is] a means for Denmark to regain its spiritual and national strength” (p. 6). Also in Italy: “Education was seen as a means of increasing Italian power and prestige” (p. 7).

The second step on the route towards mass public education was “legislation to make schooling compulsory”. Although most states had some legislation of education in place in the 18th century, often it was neither mandatory, nor universal, nor free. Secondly, in some cases legislation was set up but the necessary funds were not delivered and thus the legislation was not put into practice, as was for instance the case in Prussia in 1717 and in Denmark in 1739. Prussia was one of the first regions which provided state mass education and also established a tax instrument to finance it, but only after 1817.

With the legislation in place, the “creation of a state educational ministry or department” and the “establishment of state authority over existing and new schools” were possible. Prussia established a Bureau of Education in 1806 and after 1817 the state provided certification of teachers, while in England only in 1944 a national ministry of education was formed to act as the central education authority. Churches often had influence on education until well in the 19th century. In Austria in 1855 full control of education was given back from the state to the church and even in Prussia, the government announced in 1844 that “religious instruction is more important than pedagogical theory”.

Before the state took control of the (existing and new) schools, schools with different curricula and structures could exist side by side. This provides an explanation to the

current existence of differences in differentiation and the different levels of standardization between countries. When different school types exist next to each other and they have to merge into one education system, it might be chosen to keep the different school systems as complements, i.e. as different tracks. And it might also be that when these school merge into one system, the schools are allowed to keep some aspects of their former system (for instance their pedagogical views and related textbooks and courses) and therefore standardization might be low. How different countries came to different solutions for merging multiple school types into one is not the purpose of this chapter. However, in the sociological literature a number of theories exist to explain this. For instance, Archer (1979) or Boli *et al.* (1985).

2.3 Empirical support for the influence of Napoleon on education systems

This section builds upon the previous section in the sense that it will supplement the theory by Ramirez and Boli (1987) by empirical evidence that the political pressure from Napoleon influenced specific characteristics of European education systems and that this influence is still seen today.

2.3.1 Data

The data used in this section consists of data on input standardization from the Program for International Student Assessment (PISA) 2006 organized by the Organization for Economic Co-operation and Development (OECD), data on tracking and output standardization and data on the political pressure from Napoleon. The last two are obtained from various sources.

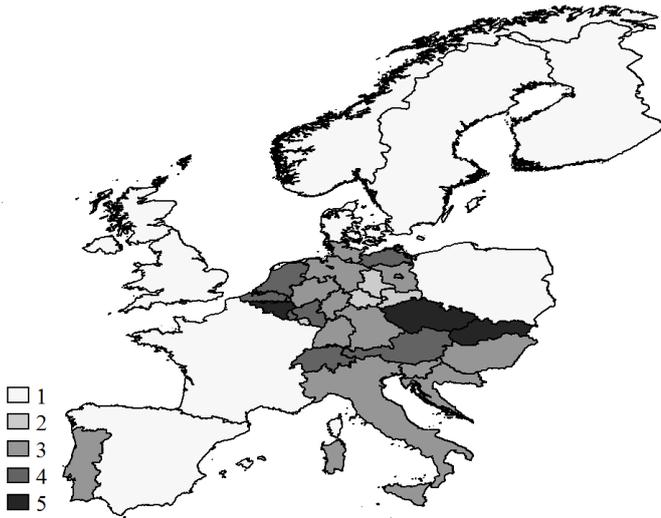
Figure 1 shows how education systems differ on two aspects: tracking and standardization. Tracking, or differentiation, can be shown by the number of tracks available to pupils and the age of first selection in the system. Data on the number of tracks is from the OECD (2007, Table 5.2) and adjusted for Flemish and French Belgium (Eurydice, 2013) and for the Germany states (Woessmann, 2007). Data on the age of first selection is from the OECD (2007, Table 5.2) and adjusted for Flemish and French Belgium (Eurydice, 2013) and for the Germany states (Kultusminister Konferenz, 2013). Table 1 and Figure 1 show the levels of tracking for 29 countries. Eight countries have one track available to students, while five tracks are available in 3 countries. The earliest selection is at age 10, while the latest is at age 16.

Figure 1: Education systems in Europe

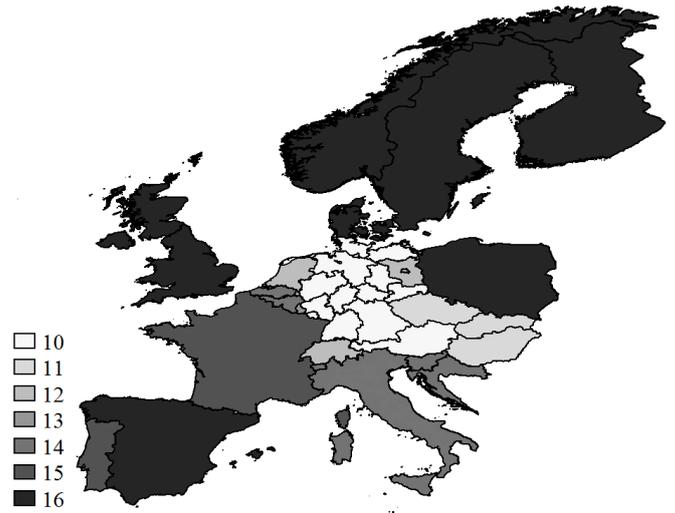
a. Number of tracks

b. Age of first selection

The Effect of Napoleon Bonaparte on Education Systems and its Consequences



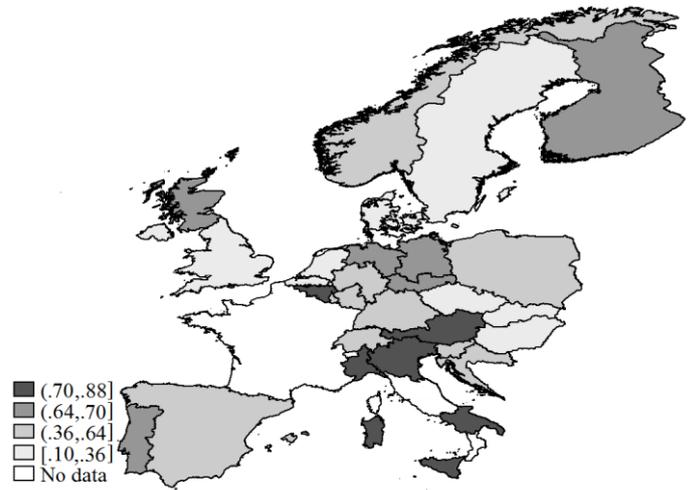
c. Central exit exams

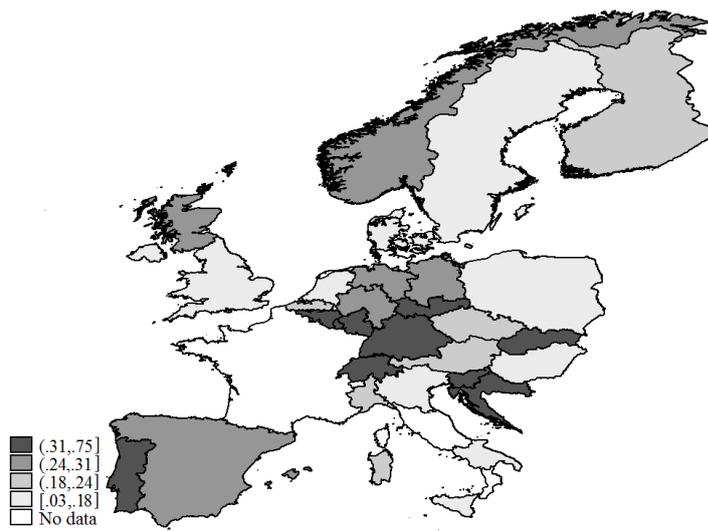


d. Administrative administration



e. Content administration





Notes: The borders are drawn at the lowest possible level at which the data are available. In the case of the number of tracks, the age of first selection and the central exit exams this means at the country level, or at the region level for French Belgium and Flemish Belgium and for England and Scotland, or at the state level for Germany. The borders for standardization deviate from this, in the sense that the lowest possible level for Germany is the state cluster level as explained in Appendix B.

Table 1: Descriptive statistics education systems in Europe⁶

| | Obs | Mean | Std. Dev. | Min | Max |
|------------------------|-----|--------|-----------|--------|---------|
| Number of tracks | 29 | 2.83 | 1.28 | 1.00 | 5.00 |
| Age of first selection | 29 | 13.07 | 2.28 | 10.00 | 16.00 |
| Central Exit Exams | 26 | 0.42 | 0.48 | 0.00 | 1.00 |
| Administrative stand. | 29 | 0.52 | 0.23 | 0.10 | 0.88 |
| Content related stand. | 29 | 0.28 | 0.16 | 0.03 | 0.75 |
| Distance in km | 29 | 940.69 | 505.55 | 264.00 | 2304.00 |
| Empire | 29 | 0.45 | 0.51 | 0.00 | 1.00 |
| Pays réunis | 29 | 0.38 | 0.49 | 0.00 | 1.00 |
| Pays conquis | 29 | 0.24 | 0.44 | 0.00 | 1.00 |
| Pays allies | 29 | 0.21 | 0.41 | 0.00 | 1.00 |

Notes: The sources of the variables are described in the text.

Standardization here is divided into two parts: standardization on output, shown by the existence of a central exit examination, and standardization on inputs, which is divided into standardization on administrative aspects and standardization on content related issues. Data on central exit examinations is from OECD (2011, Table D5.1a) and adjusted for the German states using Juerges *et al.* (2005). Data on input

6. The countries/regions used in this Table are Austria, Belgium (split into Flemish Belgium and French Belgium), Switzerland, Czech Republic, Germany (split into Bavaria, Hanover, Prussia, Rhineland, Saxony, Westphalia), Denmark, Spain, Finland, United Kingdom (split into England-Wales and Scotland), Croatia, Hungary, Italy (split into the north west, the Republic of Italy, and the Kingdom of Naples), Luxembourg, Netherlands, Norway, Poland, Portugal, Slovak Republic, Slovenia, and Sweden. See Appendix B for a more extensive description of the regions.

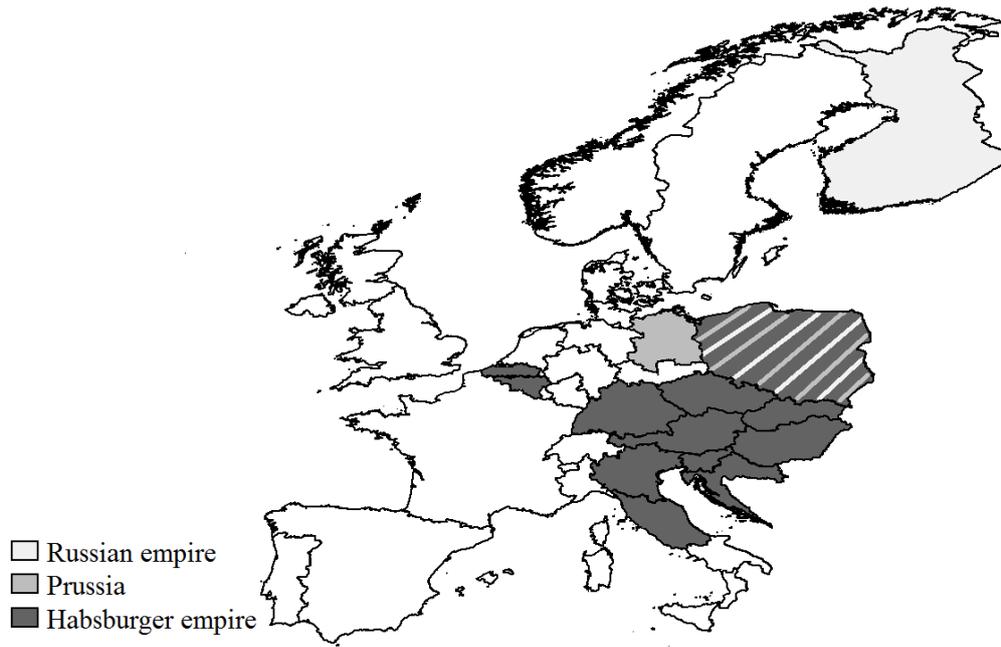
standardization is from the Program for International Student Assessment (PISA) 2006, which included answers given by school principals on who has authority for the textbooks, course content, teacher appointments, teacher employment conditions, and budget. These last five variables are transformed into two country averages, one on the administrative aspects (who is responsible for teacher hiring, salary increases and formulating the budget) and standardization on content related issues (who is responsible for textbook choice and course content). Figure 1 shows that countries in the periphery of Europe have less differentiation (both measured by the age of first selection and by the number of tracks) than countries in the inner region, but for standardization the picture is less clear.

The geographical spread in Figure 1 shows a (weak) relation between France and education systems. This geographical spread is best captured by modeling the political threat from Napoleon by distance of one's own capital to Paris. The air distance between Paris and the capital around 1800 in Table 1 is given in kilometers, 500 kilometers is added for a sea crossing. Belgium's Brussels is closest to Paris with a distance of 264 km, while Finland's capital of Turku is furthest away with 2304 km. Distance to Paris is a measure of the ex-ante political threat by Napoleon, since it uses the theoretical threat.

Whether or not countries belonged to an empire in the 1800s controls for any protection or incitement this might have ensured. Belonging to an empire would have either protected the country from outside threat due to the larger protective force of an empire compared to a (smaller) country and thus lowered the threat from Napoleon, or it would have increased the threat when there were struggles between Napoleon and the empire in question. The empires in Europe in the 1800s were the Russian Empire, the Habsburg Empire, and Prussia, as shown in Figure 2.⁷ Together with the distance to Paris, this provides a model to measure the ex-ante political threat of Napoleon.

7. Prussia was a Kingdom and not an empire but it will be referred to one in this chapter for simplicity. A fourth empire in Europe was the Ottoman Empire. However, only one country (Croatia) that once belonged to the Ottoman Empire is included in this chapter.

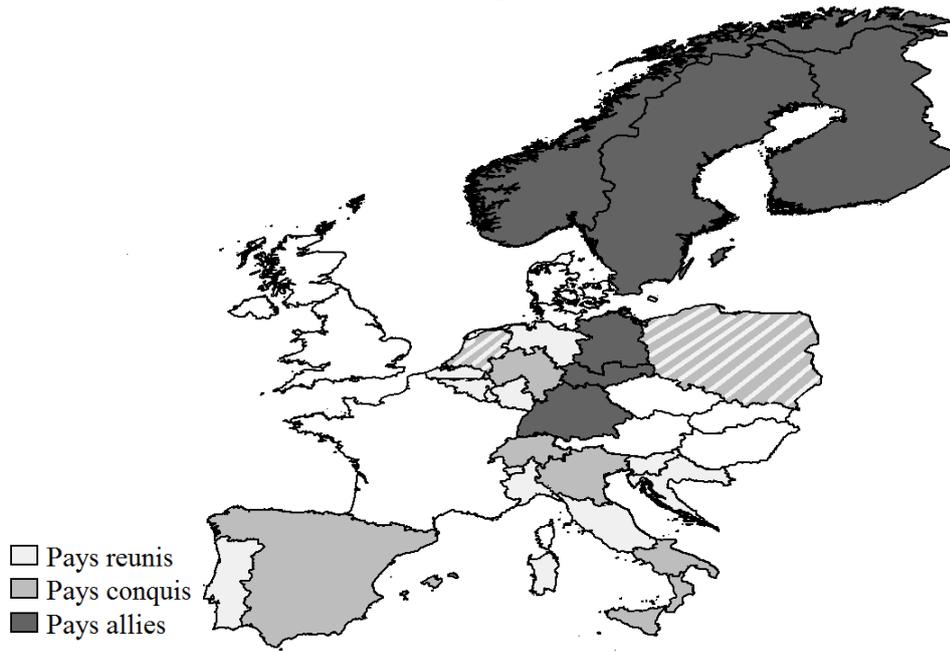
Figure 2: Empires in Europe around 1800s



Notes: These maps are based on the Centennia Historical Atlas. Appendix B provides a justification of the borders used in this chapter. Regions of Poland belonged to either Prussia (the north west), the Habsburger Empire (the south), and the Russian Empire (the east).

A measure of the ex-post threat used in this chapter is the categorization by Grab (2003), as shown before. This is ex-post since it relates to which countries were actually annexed and not which countries were under threat. The categorization from Grab (2003) divides the European countries into “countries annexed to France and directly ruled by Napoleon (*pays réunis*)”, conquered countries which were “satellite states that were entrusted to French rulers (*pays conquis*)”, and “allied countries whose territory was expanded by Napoleon in some cases, and continued to be governed by their native rulers (*pays allies*)” and is shown in Figure 3. Descriptive statistics are shown in Table 1. Appendix C provides more information on the categorization.

Figure 3: The reach of the French Empire (based on Grab, 2003)



Notes: *Pays réunis* are “countries annexed to France and directly ruled by Napoleon”. *Pays conquis* are conquered countries, or “satellite states that were entrusted to French rulers”. *Pays allies* are “allied countries whose territory was expanded by Napoleon in some cases, and continued to be governed by their native rulers”. The Netherlands and Poland were both *pays réunis* and *pays conquis*. Appendix B provides a justification of the borders used in this chapter. Appendix C provides the full list of the categorization of regions and countries by Grab (2003).

2.3.2 Results

Table 2 shows different models relating the distance to Paris, whether or not countries belonged to an empire, and the categorization by Grab (2003) to the number of tracks available in a country. The first model shows only the relation between distance and the number of tracks available and shows a strong negative correlation, meaning that the further away one’s own capital is from Paris the fewer tracks its education systems will have. Model 2 uses both the distance and the empire dummy, and is therefore the preferred model since it relates most closely to the ex-ante political threat by Napoleon. This model has a lot of power, explaining forty-five percent of the variation in differentiation. Countries belonging to an empire in 1800 have more tracks available to students than countries not belonging to an empire in 1800. The coefficient for the distance to Paris is similar to the previous model.

Model 3 and 4 use the historical categorization of Grab, both combined with distance and alone. The fourth model with only the three Grab dummies seems less able to explain the variation found in the number of tracks in Europe. None of the dummies are statistically significantly from zero, but an F test shows that jointly they are. Model 5 includes all independent variables and also these are jointly significant. A possible reason that the variables in model 3, 4 and 5 are jointly significance but that the Grab dummies are not significant is the high correlation between the three: Only seven

countries are not *pays réunis*, nor *pays conquis*, nor *pays allies*.⁸ All other countries are *pays réunis*, or *pays conquis*, or *pays allies*, except for Poland and the Netherlands which are both *pays réunis* and *pays conquis* since their status changed over time. The seven countries which are neither *pays réunis*, nor *pays conquis*, nor *pays allies* are therefore the reference category for the categorization by Grab (2003).

The five models do remarkably well: They explain fifteen to forty-five percent of the variation and the variables in all models are jointly significant. The estimates say that countries that have their capital further away from Paris have fewer tracks available to pupils nowadays. Countries that were part of an empire have more tracks. Annexed countries ruled by Napoleon (*pays réunis*) have more tracks, while other annexed or allied countries have fewer tracks. However, the estimates for the Grab dummies might be difficult to interpret due to the correlation between the three.

Table 2: The effect of political pressure from Napoleon on the number of tracks

| | (1) | (2) | (3) | (4) | (5) |
|-------------------------|----------------------|----------------------|---------------------|---------------------|---------------------|
| Distance | -0.002*** (0.000) | -0.002*** (0.000) | -0.001** (0.001) | | -0.001** (0.000) |
| Empire | | 1.031** (0.375) | | | 1.094*** (0.387) |
| Pays réunis | | | 0.486 (0.532) | 0.992* (0.540) | 0.726 (0.476) |
| Pays conquis | | | -0.397 (0.537) | -0.261 (0.586) | -0.097 (0.484) |
| Pays allies | | | -0.151 (0.649) | -0.692 (0.672) | -0.051 (0.573) |
| Constant | 4.344*** (0.444) | 3.932*** (0.426) | 4.053*** (0.677) | 2.692*** (0.436) | 3.326*** (0.649) |
| # of countries | 29 | 29 | 29 | 29 | 29 |
| Adjusted R ² | 0.32 | 0.45 | 0.29 | 0.15 | 0.45 |
| p F test | 0.001 | 0.000 | 0.014 | 0.073 | 0.002 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and the number of tracks in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See table 1 for notes on the data sources. Model 2 is the preferred model.

Table 3 shows the same models but now for the age of first selection. Also here the first three models are capable of explaining part of the variation in the age of selection, but the fourth model is not, as seen by the F test and the adjusted R². The coefficients have the reversed signs as compared to Table 2, as is to be expected since the number of tracks available in a country and the age of first selection are highly negatively correlated (-0.66 in this sample).

8. These are Austria, Czech Republic, Denmark, England, Hungary, Scotland, and Slovak Republic.

Table 3: The effect of political pressure from Napoleon on the age of first selection

| | (1) | (2) | (3) | (4) | (5) |
|-------------------------|----------------------|----------------------|---------------------|----------------------|----------------------|
| Distance | 0.002** (0.001) | 0.002** (0.001) | 0.003*** (0.001) | | 0.003*** (0.001) |
| Empire | | -0.919 (0.807) | | | -0.760 (0.828) |
| Pays réunis | | | 0.788 (0.998) | -0.451 (1.084) | 0.621 (1.018) |
| Pays conquis | | | 0.574 (1.007) | 0.241 (1.177) | 0.365 (1.035) |
| Pays allies | | | -1.643 (1.218) | -0.316 (1.348) | -1.712 (1.224) |
| Constant | 11.140*** (0.863) | 11.510*** (0.917) | 9.982*** (1.270) | 13.320*** (0.875) | 10.490*** (1.387) |
| # of countries | 29 | 29 | 29 | 29 | 29 |
| Adjusted R ² | 0.17 | 0.18 | 0.20 | -0.11 | 0.19 |
| p F test | 0.015 | 0.028 | 0.053 | 0.96 | 0.074 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and the age of first selection in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See Table 1 for notes on the data sources. Model 2 is the preferred model.

Table 4 shows the models using central examinations as dependent variable. What is immediately clear is that although the pressure of Napoleon has a relation to the two measures of differentiation, it has no such relation with central exit examinations. No model has jointly significant variables. Perhaps the existence of central exit examination is too recent to be affected by the political pressure from Napoleon, or the use of central exit examinations changed too much over the years.

Table 5 shows how that the pressure from Napoleon is also not related to the level of standardization on administrative issues. None of the coefficients is significant, also not jointly.

Table 4: The effect of political pressure from Napoleon on the existence of central examinations

| | (1) | (2) | (3) | (4) | (5) |
|-------------------------|-------------------|--------------------|-------------------|------------------|-------------------|
| Distance | 0.000 (0.000) | 0.000 (0.000) | -0.000 (0.000) | | -0.000 (0.000) |
| Empire | | -0.170 (0.207) | | | -0.180 (0.215) |
| Pays réunis | | | 0.031 (0.269) | 0.077 (0.240) | -0.039 (0.284) |
| Pays conquis | | | 0.401 (0.262) | 0.407 (0.257) | 0.364 (0.268) |
| Pays allies | | | 0.463 (0.297) | 0.419 (0.272) | 0.463 (0.299) |
| Constant | 0.414* (0.214) | 0.484** (0.232) | 0.350 (0.310) | 0.247 (0.180) | 0.489 (0.354) |
| # of countries | 26 | 26 | 26 | 26 | 26 |
| Adjusted R ² | -0.04 | -0.05 | 0.00 | 0.03 | -0.02 |
| p F test | 0.802 | 0.695 | 0.445 | 0.304 | 0.498 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and the existence of central exit exams in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See Table 1 for notes on the data sources. Model 2 is the preferred model.

Table 5: The effect of political pressure from Napoleon on the administrative standardization

| | (1) | (2) | (3) | (4) | (5) |
|-------------------------|----------------------|---------------------|---------------------|----------------------|---------------------|
| Distance | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | | 0.000 (0.000) |
| Empire | | 0.005 (0.089) | | | 0.021 (0.096) |
| Pays réunis | | | 0.040 (0.114) | 0.053 (0.104) | 0.045 (0.119) |
| Pays conquis | | | 0.081 (0.115) | 0.084 (0.113) | 0.087 (0.121) |
| Pays allies | | | 0.085 (0.139) | 0.072 (0.129) | 0.087 (0.143) |
| Constant | 0.547*** (0.0930) | 0.545*** (0.101) | 0.500*** (0.145) | 0.466*** (0.0837) | 0.486*** (0.162) |
| # of countries | 29 | 29 | 29 | 29 | 29 |
| Adjusted R ² | -0.03 | -0.07 | -0.13 | -0.09 | -0.18 |
| p F test | 0.758 | 0.953 | 0.941 | 0.871 | 0.976 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and administrative standardization in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See Table 1 for notes on the data sources. Model 2 is the preferred model.

The relation between the pressure from Napoleon and the second form of input standardization, content related standardization, is very small, but model 3 and 4 of

Table 6 do show jointly significant correlations. Here especially, annexed countries rules by Napoleon (*pays réunis*) are seen to have more content related standardization.

Table 6: The effect of political pressure from Napoleon on the content related standardization

| | (1) | (2) | (3) | (4) | (5) |
|-------------------------|---------------------|---------------------|---------------------|---------------------|--------------------|
| Distance | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | | 0.000 (0.000) |
| Empire | | 0.011 (0.060) | | | 0.016 (0.058) |
| Pays réunis | | | 0.140* (0.069) | 0.164** (0.0636) | 0.144* (0.072) |
| Pays conquis | | | -0.052 (0.070) | -0.0460 (0.0690) | -0.048 (0.073) |
| Pays allies | | | 0.095 (0.084) | 0.0697 (0.0791) | 0.097 (0.086) |
| Constant | 0.357*** (0.063) | 0.353*** (0.069) | 0.276*** (0.088) | 0.213*** (0.051) | 0.266** (0.098) |
| # of countries | 29 | 29 | 29 | 29 | 29 |
| Adjusted R ² | 0.04 | 0.00 | 0.16 | 0.17 | 0.13 |
| p F test | 0.166 | 0.384 | 0.0865 | 0.0581 | 0.153 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and content administration in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See Table 1 for notes on the data sources. Model 2 is the preferred model.

All in all, pressure from Napoleon does seem to have a relation with some aspects of education systems in Europe; most notably with differentiation, as seen by the number of tracks and age of selection, and content related standardization. For the number of tracks in a country, the proposed variables are able to explain thirty to forty-five percent of the cross country variation. For the age of selection the pressure from Napoleon is able to explain around twenty percent. These results therefore provide support for the thesis of Ramirez and Boli (1987).

2.4 Can the Napoleonic Wars explain student performance?

Section 2.3 shows that the political pressure from Napoleon has a relation with some current characteristics of education systems in Europe. Therefore, it is possible to investigate the effect of education systems characteristics on student performance using pressure from Napoleon. In general, studying the effects of education systems characteristics on student performance is hindered by the endogeneity of education systems, in which factors that shaped the current characteristics also (direct or indirect) influence current student performance. Using Instrumental Variable (IV) analysis this endogeneity can be resolved, and in this chapter pressure from Napoleon is used as an

instrument for education systems characteristics. Section 2.5 pays closer attention to the exclusion restriction.

This chapter looks at two aspects of education systems: tracking and standardization. From Section 2.3 it is to be expected that the IV will work best for tracking since there the model could explain the cross country variation in system characteristics. For standardization, especially central exams, less is to be expected. Therefore the following section focuses on tracking. Standardization is not pursued any further, but Appendix D does present the results for the interested reader.

2.4.1 Data

Data on three topics are used. Individual level data on student performance and background characteristics are taken from PISA 2006. Country level data on tracking and on the political threat from Napoleon are as discussed in Section 2.3.

Student performance is obtained from PISA 2006 which was conducted in 57 countries, of which 21 are in Europe and located on the battle ground of the Napoleonic Wars. France is excluded from both samples since the argument of political threat from France does not apply to France itself. Four countries are split into regions: Belgium (into Flemish and French Belgium), the United Kingdom (into England/Wales and Scotland), Italy (into the North West, the Republic of Italy and the Kingdom of Naples⁹), and Germany (into Bavaria, Hanover, Prussia, Rhineland, Saxony, Westphalia).¹⁰ This chapter therefore uses 29 regions, which will be called countries for simplicity, 6578 schools, and 157,200 students.¹¹ Missing values are replaced by either school or country averages. Analyses are done with student weights and clustering of errors on country level. Using country level variables on differentiation and clustering on the country level implies that the sample effectively only contains 29 observations.

PISA 2006 contains a reading, mathematics, and science test score. This chapter uses the mathematics test score. However, the results are very similar when I use the science test score, which was the main focus of PISA 2006, or the reading test score, although the reading results are not significant. Next to the test scores, PISA also provides information on the student and the school. In this chapter the following control variables are used: the gender of the student, the number of books in the household (as a proxy for parental background), the grade of the student, whether the student is in vocational education, the school average of the number of books of the

9. Some Italian regions are excluded since in PISA they are grouped together in a non-geographical way.

10. For more information on these regions, see Appendix B.

11. Excluded are students with no information on age, gender or social economic background, migrant students, and students in schools with less than 5 students in PISA.

students, the school size, the location of the school, whether the school has ability grouping, and GDP per capita 2005.

Table 8: Correlations between age of first selection and test scores

| Database: Subject: | PISA 2000 reading | PISA 2006 reading | PISA 2006 math | PISA 2006 science |
|---|----------------------|----------------------|-------------------|----------------------|
| <i>Uses the OECD (2010) classification based upon age of first selection</i> | | | | |
| OECD (2010) sample (27 cnts) | 0.59 | 0.34 | 0.23 | 0.26 |
| OECD (2010) minus MEX | 0.54 | 0.23 | 0.07 | 0.10 |
| Including NLD, HRV, SVK, SVN | 0.43 ^a | 0.25 | 0.04 | 0.09 |
| Sample used in this chapter (20 cnts) | 0.37 ^b | 0.21 | -0.04 | 0.02 |
| <i>Uses age of first selection from OECD (2007)</i> | | | | |
| Sample used in this chapter (20 cnts) | 0.19 ^b | 0.19 | -0.16 | -0.11 |
| Sample used in this chapter (29 regions) | - | -0.14 | -0.23 | -0.35 |
| <i>Notes:</i> ^a In PISA 2000 only the Netherlands can be added. ^b This sample misses Croatia, Slovak Republic and Slovenia. | | | | |

The same two aspects of differentiation in education systems are used here as in Section 2.3: the number of tracks available to 15-year-old pupils and the age of first selection. The correlation between the mathematics test score and the age of first selection is -0.23, while the correlation with the number of tracks at the country level is 0.06. These correlations deviate widely from the consensus in the public debate as illustrated by the following correlations reported by the OECD (2010): the country level correlation between reading scores and a classification based on the age of first selection in PISA 2000 is 0.59 for the sample of OECD countries (OECD, 2010, p. 58). This difference in correlations between this chapter and OECD (2010) is unraveled in Table 8. First, the OECD correlation is based on PISA 2000, while in this chapter PISA 2006 is used (making the correlation drop from 0.59 to 0.34). Second, the OECD used the reading test score, while I use the mathematics test score (0.34 vs 0.23). Third, I add five countries (the Netherlands, Croatia, Norway, Slovak Republic, and Slovenia) and drop all non-European countries (Australia, Canada, Japan, Korea, Mexico, New Zealand, and the United States) and some European countries which were not in the sphere of influence of France (France itself and Greece, Iceland and Ireland) (0.26 vs -0.04). However, the country that most drives the high positive correlation between age of first selection and the reading score in PISA 2006 is Mexico, as can be seen in the second row where Mexico is excluded from the OECD (2010) sample. Fourth, I split Belgium, Germany, Italy, and the United Kingdom into regions and in the analyses I use the age of first selection variable as provided by the OECD (2007), while the OECD (2010) uses a classification based upon the age of first

selection.¹² This results in a correlation between the age of first selection and student performance of -0.23.

2.4.2 Estimation Method

The models estimated in this paper are straightforward IV models shown by equations (1) and (2). In this first stage, I use the political pressure from Napoleon ($NAPOLEON_c$) to capture the variation in tracking ($TRACKING_c$) due to this pressure for each country c . In the second stage, I then use the fitted values to estimate the effect of tracking on student performance ($TEST SCORE_{i,c}$). I control for individual, school and country characteristics captured in $X_{i,c}$. These background characteristics are the gender of the student, the number of books in the household, the grade of the student compared to the modal grade in the country, whether the student is in vocational education, the school average of the number of books of the students, the school size, the location of the school, whether the school has ability grouping, and GDP per capita in 2005.¹³ The variable of interest is β which displays the effect of tracking on student performance.

$$TRACKING_c = \gamma + \delta NAPOLEON_c + X_{i,c}\theta + \epsilon_i \quad (1)$$

$$TEST SCORE_{i,c} = \alpha + \beta \widehat{TRACKING}_c + X_{i,c}\vartheta + \epsilon_i \quad (2)$$

For a valid IV two conditions must hold: First, the instrument should have an effect on the endogenous variable. Second, the instrument should not have an effect on anything else that might influence the outcome, in this case, student performance. The second condition (the exclusion restriction) cannot be tested and must be taken on good faith. Section 2.5 elaborates more on this. The first condition can be verified by a strong first stage. Sections 2.3.2 and 2.4.3 show that this first condition is indeed fulfilled.

In these analyses the level of tracking is captured by two variables: the number of tracks available to 15 year old students and the age of first selection. In general, using the same instrument for multiple endogenous variables is problematic since this will violate the exclusion restriction. However, both the number of tracks and the age of first selection are proxies of the underlying level of tracking in an education system justifying the use of a single theory behind the instruments for both. The consequence is that the results have to be interpreted either separately for both endogenous variables by assuming no effect of the other aspect of tracking or jointly. A third option would

12. Countries with no selection before or at 15 are group 1, countries with selection at age 14 or 15 are group 2 and countries with selection before 14 are group 3. Between the age of first selection used in OECD (2010) and the variable used in this chapter are two differences: Switzerland selects at age 14 (instead of age 12) and Poland selects at age 16 instead of 15.

13. Missing values are replaced by school or country averages and imputation dummies and interactions are included to ensure the imputation method does not affect the estimates.

be to use more instruments for both aspects simultaneously. In the discussion of the results, this is further addressed.

2.4.3 Results

The negative correlation between the age of first selection and student performance shown in Table 8 is most probably clouded by endogeneity. The current section tries to alleviate this endogeneity by instrumenting differentiation by the political pressure from Napoleon as discussed in Section 2.3. Table 9 shows the OLS and IV results of differentiation on student performance, using the same model specifications as before.¹⁴ The upper panel of Table 9 shows the first stage, while the lower panel shows the second stage. In Section 2.3, it was shown that it was most promising to use the political pressure from Napoleon as an instrument for differentiation and also here the first stages are quite strong and very similar to in Section 2.3. Using the first dimension of differentiation, the number of tracks, reveals a significant effect of the number of tracks on student performance as seen by mathematics test scores. The OLS model reveals that one more track available to students leads to an 8.8 point (**) increase in student math test scores. Model 2 and 4 show that this estimate is downwards biased since the IV models reveal a coefficient of 16.3-16.7** of the number of tracks. The point estimates of the other models are smaller (11.8-13.0) and the standard errors are bigger, resulting in insignificant coefficients. The downward bias in the OLS estimates could be explained by the endogeneity in the data for which I correct using IV. It must be that some countries which had bad student performance moved towards more tracking after the influence of Napoleon stopped or countries that had good performance shifted away from tracking, which explains the lower difference in the average student performance found by OLS.

The preferred model 2 says that when a system has one more track, student performance as measured by a math test score will be one-sixth of a standard deviation higher (mean: 507; sd: 93.1). This seems small, but it is 1.5 times the amount of moving an individual student from a household with 0 to 10 books to a household with 11 to 25 books (10.91***) which is quite substantial.¹⁵ The full second stage model including all the control variables is depicted in Appendix E.

The second dimension of tracking, the age of first selection, has no significant effects on student performance, as shown by the right panel of Table 3, although all estimates are consistently negative. However, note that the two measures of tracking are correlated and thus the results for both measures must either be interpreted separately by assuming the insignificance of the other or jointly. A joint interpretation would be that the level of tracking has a positive effect on student performance, but that it is unclear whether this comes from a positive effect of the number of tracks or a negative effect of the age of first selection. IV models using more than two instruments for both endogenous variables simultaneously resulted into similar coefficients (not shown, but available on request). However, the standard errors were much larger leading to

14. Model 4 is excluded since in Section 2.2 it was the weakest model.

15. In 2006 the variable *number of books in the household* has six categories: 0-10 books, 11-25 books, 26-100, 101-200, 201-500, more than 500.

insignificant results. This indicates that with only 29 countries this simultaneous estimation is too demanding for the data.

It is interesting to relate these results to those found by others. The literature on the effect of tracking is not as extensive, due to the problem of endogeneity, and the findings are not very consistent. Hanushek and Woessmann (2006) provide a difference-in-difference analyses into the effects of selection into tracks at an early age and find a negative effect of early tracking (although not robust to sample changes), which is confirmed by Van Elk *et al.* (2011). Ariga and Brunello (2007) look at the tracking length and find a positive effect of the time spend in tracks. Fuchs and Woessmann (2007), Pekkarinen (2008), and Kerr *et al.* (2013) find insignificant effects. The OECD correlations reported earlier are therefore neither widely confirmed by the literature, nor by this study.

Table 9: The effect of differentiation on mathematics test scores

| | Number of tracks | | | | | Age of first selection | | | | |
|---|--------------------|----------------------|----------------------|----------------------|----------------------|------------------------|---------------------|---------------------|---------------------|---------------------|
| | (0) | (1) | (2) | (3) | (4) | (0) | (1) | (2) | (3) | (4) |
| 1st stage | | | | | | | | | | |
| Distance | | -0.001*** (0.000) | -0.001*** (0.000) | -0.001*** (0.000) | -0.001*** (0.000) | | 0.002*** (0.000) | 0.002*** (0.000) | 0.003*** (0.001) | 0.003*** (0.001) |
| Empire | | | 0.836*** (0.286) | | 0.950*** (0.280) | | | -0.609 (0.740) | | 0.0330 (0.712) |
| Pays réunis | | | | 0.455 (0.411) | 0.612 (0.382) | | | | 0.671 (0.813) | 0.676 (0.812) |
| Pays conquis | | | | -0.312 (0.432) | -0.071 (0.445) | | | | 0.536 (0.795) | 0.545 (0.814) |
| Pays allies | | | | 0.0241 (0.456) | -0.225 (0.343) | | | | -1.967** (0.861) | -1.976** (0.870) |
| 2nd stage | | | | | | | | | | |
| Number of tracks | 8.831** (3.403) | 11.810 (9.434) | 16.330** (7.630) | 12.960 (9.180) | 16.730** (7.435) | | | | | |
| Age of first selection | | | | | | -1.991 (1.875) | -7.754 (6.901) | -10.01 (6.858) | -6.066 (4.522) | -5.974 (4.506) |
| # of students | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 |
| # of countries | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 |
| R ² 2 nd stage | 0.33 | 0.34 | 0.33 | 0.33 | 0.63 | 0.33 | 0.31 | 0.30 | 0.32 | 0.32 |
| Partial R ² 1 st stage | | 0.326 | 0.423 | 0.376 | 0.483 | | 0.210 | 0.224 | 0.338 | 0.338 |
| F test 1 st stage | | 64.89 | 27.61 | 19.34 | 15.34 | | 24.26 | 10.64 | 12.53 | 10.41 |
| <p><i>Notes:</i> The table presents coefficients (standard errors in parenthesis) on the effect of tracking on student performance, instrumenting in models (2) to (5) tracking with the political pressure from Napoleon. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Model 2 is the preferred IV model. The models include imputation dummies and imputation variable interaction terms, and clustering is done on the country level. The control variables are as explained in the text. Differences in the first stage as compared to Section 2 are due to the inclusion of the student level.</p> | | | | | | | | | | |

2.5 Threats to the identification

2.5.1 Exclusion restriction

It seems quite reasonable that the pressure from Napoleon influenced more than just tracking in education systems (e.g. Acemoglu *et al.*, 2011) and also other historical events, likely to be correlated to Napoleon, influenced parts of the human capital development of nations (e.g. Becker and Woessmann, 2009; Dittmar, 2011; Cantoni and Yuchtman, 2014). These correlated events could potentially bias the IV results. The first stage analyses in Section 2.4.3 are therefore also performed by using the GINI coefficient, the number of police officer per 100,000 inhabitants, the expenditure on schooling per capita and the number of women seats in parliament.¹⁶ Table 10 shows the effect of the political pressure from Napoleon on these outcomes, using model 2 of Table 9 (the preferred model). What can be seen is that in general political pressure from Napoleon does not really affect these current day societal factors: The individual instruments have in two cases a significant effect, but in most instances they do not. The F test on joint significance shows the model on educational spending is significant (at the 10 percent level), but for the other three models this is not the case.

Table 10: The effect of Napoleon on other societal characteristics

| | GINI | Police | Exp. Edu. | Women seats |
|-------------------------|----------------------|------------------------|---------------------|----------------------|
| Distance | -0.003* (0.001) | -0.052 (0.034) | 0.001** (0.000) | 0.003 (0.004) |
| Empire | -0.042 (1.349) | 27.040 (33.470) | -0.268 (0.378) | -0.813 (3.711) |
| Constant | 33.560*** (1.460) | 353.700*** (38.030) | 4.466*** (0.429) | 24.900*** (4.217) |
| # of countries | 26 | 29 | 29 | 29 |
| Adjusted R ² | 0.0752 | 0.03 | 0.11 | -0.05 |
| p F test | 0.156 | 0.254 | 0.087 | 0.729 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between the political pressure from Napoleon and societal characteristics. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Appendix F also shows the other models used as first stages in Table 9: The instruments using the GINI coefficient are found to be jointly significant in four out of five models, for the other societal characteristics the joint significance is only in two or one models out of five. Regardless, this suggests that at the very least pressure from Napoleon influenced (determinants of) the GINI coefficient in countries. And the level of inequality in a society, as measured by the GINI coefficient, might influence educational opportunities or other aspects of education which have an effect on student performance.

2.5.2 Changes in tracking since Napoleon

This chapter uses data from 2006 to see whether there is an effect of differentiation on student performance, using as an instrument political pressure from Napoleon in the

16. More information on this is given in Appendix D.

1800s. It could be that in-between the 1800s and current day, education systems changed, which could affect the results. It is difficult to find data on differentiation (or schooling in general) before the twentieth century, but most change in the European education systems occurred in the last decades, following after the formation of the United Nations in 1945 and the OECD in 1961. The emphasis of states on education, both for development and economic growth, increased and globalization of education and education systems started, best seen by the growing number of national reports on the state of education since that time (e.g. UNESCO, 1968a).

To take possible changes into account, Table 11 shows analyses from OLS and (the preferred) IV models using data from before most of the change. For these analyses I use the First International Science Study (FISS) from 1970 for ten European countries and construct a differentiation index using National Reports from UNESCO(1968a; 1968b; 1971a; 1971b; 1971c; 1971d; 1971e; 1973a; 1973b; 1975). These ten countries are Flemish Belgium, French Belgium, England, Finland, West-Germany, Hungary, Italy, the Netherlands, Scotland, and Sweden. More information on the data can be found in Appendix G.

The first stage results in Table 11 are very similar to before, but the second stage coefficients are much smaller in size. The left panel of Table 11 show even a negative effect of tracking in 1970, which lowers the confidence in the causal interpretation of the results of Section 2.4.3. However, these analyses use data for only ten countries, the F statistic of the first stage is below ten, and also the data quality on both student performance and tracking is questionable. For comparison reasons, the right panel shows results for a similar differentiation index for the same ten countries using PISA 2006 data. No significant effect is found using this set up and sample, although the sign is positive. Appendix G shows the results for the other model specifications used in this chapter.

Table 11: The effect of differentiation on student performance (science) in 1970 and 2006

| Model: | 1970 | | 2006 | |
|--|--------------------|---------------------|-----------------|----------------------|
| | OLS | IV | OLS | IV |
| 1st stage | | | | |
| Distance | | -0.001** (0.000) | | -0.001*** (0.000) |
| Empire | | 0.632 (0.425) | | 1.029 (0.659) |
| 2nd stage | | | | |
| Differentiation | -0.083* (0.045) | -0.251** (0.114) | 0.01 (0.062) | 0.078 (0.116) |
| # of students | 25,409 | 25,409 | 73,171 | 73,171 |
| # of countries | 10 | 10 | 10 | 10 |
| R ² 2 nd stage | 0.23 | 0.211 | 0.353 | 0.348 |
| Partial R ² 1 st stage | | 0.279 | | 0.547 |
| F test 1 st stage | | 5.219 | | 11.22 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between tracking and student performance, instrumenting differentiation with the political pressure from Napoleon. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

2.6 Conclusion

This chapter builds on the 1987 paper of Ramirez and Boli: *The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization*. In this paper, Ramirez and Boli argue that the introduction of European education systems in the nineteenth century was induced by external challenges on the continent. One of the most pronounced challenges was the political pressure from Napoleon, who dominated the continent for two decades. Almost all European countries were at one stage involved with Napoleon and many wars were fought on numerous battlefields. To unite the people and ensure a strong and willing workforce, countries introduced mass public education systems to educate the people as citizens of the nation willing to fight for independence. In this manner, political pressure from Napoleon had an influence on European education systems.

This chapter empirically tests whether the proposed mechanism of Ramirez and Boli (1987) had an effect on levels of differentiation and standardization in European education systems. Especially for differentiation, both seen by the number of tracks and age of selection, a relation between education systems and the political pressure from Napoleon in the nineteenth century is found, as theorized by Ramirez and Boli (1987). The proposed variables capturing the political pressure from Napoleon are able

to explain thirty to forty-five of the cross country variation in the number of tracks available to students and around twenty percent for the age of selection. For standardization relating to the content, school administration or the existence of central exit exams, no or a much weaker relation is found. Still the theory proposed by Ramirez and Boli (1987) is confirmed by empirical evidence on specific education system characteristics.

As a second step, the relation between political pressure from Napoleon and education systems is used in an instrumental variables analysis to analyze student performance. In general, the effects of education systems on student performance are obscured by the endogeneity of education systems, in which factors that influenced the characteristics also (direct or indirect) influence student performance. By using exogenous variation that influenced education systems but not student performance, this endogeneity can be alleviated. This chapter uses the political pressure from Napoleon as exogenous variation that can explain the level of differentiation in education systems.

Using Napoleon as an instrument for the level of differentiation in education systems leads to a strong first stage. When the number of tracks available to students is used as a differentiation measure the instruments also lead to a strong second stage. Jointly interpreting the results on the number of tracks and the age of first selection I find a positive effect of tracking on student performance. The analyses are unable to tell whether this originates from a positive effect of the number of tracks or a negative effect of the age of first selection.

One important caveat of this analysis is that political pressure from Napoleon influenced many facets of European countries. However, this chapter provides an additional approach which supplements the existing evidence on the effects of education system characteristics on student performance. From this chapter, it can be said that on some dimensions (mainly differentiation) there is a strong and robust relation between the political pressure from Napoleon and education system characteristics. Consequently, using the political pressure from Napoleon as an instrument for differentiation leads to a strong first stage. The resulting analyses show a consistent positive relation between differentiation and student performance, while it is only significant for the number of tracks available to students.

Appendices

Appendix A: Summary of the historical case studies of Ramirez and Boli (1987) (Chapter 2)

Prussia

Prussia already provided for quite some schooling in the 18th century. For instance, in 1716 Frederick William I “made attendance at village schools compulsory for all children not otherwise provided with instruction” but funds were not provided for. But after 1806, when Prussia lost to the French army and was subjected to large influence from France, Prussia made great strides in its education efforts. The German philosopher Johann Gottlieb Fichte in 1807 “claimed that universal state-directed, compulsory education would teach all Germans to be good Germans and would prepare them to play whatever role [...] fell to them in helping the state reassert Prussian power” (p. 5). In the same year the Bureau of Education was established and between 1817 and 1825 a state administration was set up and taxes to finance education were imposed.

Austria

Austria introduced its first universal compulsory school law in 1774 but the implementation was “frustrated by the reactionary policies in the aftermath of the French Revolution and the Napoleonic Wars” and education came back in religious control. In 1866 education came back in state control, but only after 1869 the first compulsory school law was accepted.

Denmark

In Denmark multiple attempts were made towards compulsory schooling in the 18th century, but money was never allocated to make the attempts successful. The final, and successful, attempt to introduce a Danish education system was between 1789 and 1814 under commission of Frederick VI. This period coincide with the fall of Denmark in the early 19th century. N.F.S. Grundvig, a clergymen, saw education “as a means for Denmark to regain its spiritual and national strength”.

Sweden

Before and during the Napoleonic War there more numerous proposals in Sweden on education, but they were either religiously inspired or they were defeated by clergy, aristocracy or king, who was a former general from Napoleon. After the Napoleonic Wars there was more room for liberal reform and a school reform bill in 1842 was adopted. “In Sweden, the crisis brought about by the Napoleonic Wars was quite acute, but it did not immediately generate a state educational system, despite efforts in that

direction [...] [due to] the resistance of the relatively powerful and independent peasantry (together with the standard resistance of the aristocracy and clergy) that slowed the movement towards state-controlled education” (p. 7).

Italy

Italy had its “first serious calls for state-directed schooling [...] in 1796 when Austria’s counterattack on France seemed likely to quench the [French] Revolution” (p.7). But little was achieved. Only after Italian unification was real progress made on establishing a mass public education system. However, “The Piedmontese state strove to build a national education system before national unification had been achieved” (p. 7). It was thus the north-western state of Piedmont, closest to France, that was first to seriously consider mass education in Italy.

France

“In France, education under the *ancient regime* consisted mainly of secondary schools run by religious orders for the middle and upper class and the school runs by Catholic Church”. Although Napoleon frequently pointed towards the importance of education, he ignored primary education and focused purely on elite secondary schooling (*lycées* and *Grandes Ecoles*). Only after 1833 did France made the first efforts towards mass primary education, but these efforts failed due to the “Revolution of 1840 and the regime of Louis Napoleon in 1852” (p.8). Between 1870 and 1881 an elementary school system was established.

England

In England, education was restricted to the elite and the “the movement towards state-sponsored mass education was inhibited by the very success of England’s navy and merchants [...]. Schooling developed slowly in private hands and classical liberal restraints on state action kept the state out of education much longer in England than elsewhere.” (p. 8). The first successful reform was in 1833 and only stipulated state grants for schools. After 1860 the situation changed and debate also shifted to providing education for the working classes. In 1870 the Elementary Education Act was established which provided mandatory primary education. Church control of schools was ended in 1902.

Appendix B: The country borders (Chapter 2)

The borders of the maps used in this chapter are drawn by using the command `spmap` in Stata making use of the Shape files and the dBase databases of the European countries found on www.gamd.org. Thus the maps here are drawn using the current borders of the European countries. This means that country borders in the historical

maps could (and most probably will) deviate from the historical borders in a number of ways. For simplicity it is chosen here to ignore these changes.

In a few instances was it possible to divide countries into regions more suitable for either the historical categorization of the French treat (Germany and Italy) or for the display of education systems (Germany, Belgium and Great Britain).

Germany and Italy

Germany and Italy were in the 1800s not yet the united countries as they are today. In Germany a large number of states, city states and kingdoms existed and the borders and the independency of these regions changed numerous times. To be able to match the current German states to the regions in the 1800s a division of Germany is made. In this chapter Westphalia refers to the current German states Hesse and North Rhine-Westphalia. The Rhineland are the German states Saarland and Rhineland-Palatinate. Prussia consists of Mecklenburg-Vorpommern, Berlin, Brandenburg, and Saxony-Anhalt. Bavaria is Bavaria and Baden-Wurttemberg. Hanover consists of the German states Hamburg, Bremen, and Lower Saxony, while Saxony refers to Saxony and Thuringia.¹⁷ Schleswig-Holstein belonged to Denmark around 1800 and is therefore display as a single territory in the historical maps and is excluded from the maps depicting standardization. The standardization maps are made using PISA2006 data and it would be very difficult and to some extent arbitrary to somehow add the Danish and Schleswig-Holstein data to create one “Greater Denmark”.

Italy was similar to Germany in the 1800s in the sense that it consisted of a number of states and kingdoms and was only later united in a single country. This chapter distinguishes four regions in Italy in the 1800s: The North West consisting of current day Liguria, Piedmont, and Sardinia, The Republic of Italy consisting of Bolzano, Trento, Emilia-Romagna, Friuli-Venezia Giulia, Lombardy, and Veneto, The Kingdom of Naples: Basilicata, Campania, Puglia, Provincia Sicilia. The other Italian provinces are Abruzzo, Lazio, Marche, Molise, Tuscany, Umbria, Valle d'Aosta, and Calabria. Of these eight provinces the first six can be regarded as the Papal State while Valle d'Aosta could be assigned to the North West and Calabria was part of the Kingdom of Naples. However, in PISA2006 it not possible to distinguish between these eight provinces and therefore these provinces are excluded from the maps and analysis using standardization.

17. The division of German states into these regions, instead of using all German states separately, is used since it is not allowed to display results using PISA2006 with individual German states. It is allowed however to use clusters of states, which is what is done in this chapter.

Belgium and Great Britain

Both Belgium and Great Britain are divided in two regions that better represent the education systems (and also the cultural regions). Belgium is divided in the Flemish region and the French region (including the German speaking parts), while Great Britain is divided in Scotland and the region consisting of England, Wales, and Northern-Ireland.

Appendix C: Historical categorization of the French influence by Grab (p. 18, 2003) (Chapter 2)

Pays réunis: “countries annexed to France and directly ruled by Napoleon”.

Avignon (1791)

Savoy (1792)

Nice (1793)

Belgium (1795)

Luxemburg (1795)

Geneva (1798)

Piedmont (1802)

The Rhineland (1802)

Liguria (1805)

Kingdom of Etruria (Tuscany) (1808)

Parma (1808)

Rome and its environs (1809)

The Illyrian Provinces (1809)

Kingdom of Holland (1810)

Hanover and Hamsa cities of Hamburg, Bremen and Lubeck (1810)

The Grand Duchy of Oldenburg (1810)

Pays conquis: conquered countries, or “satellite states that were entrusted to French rulers”.

Swiss Confederation (created in 1803)

Republic of Italy (1802)/Kingdom of Italy (1805)

Kingdom of Naples (1806)

Grand Duchy of Berg (1806)

Kingdom of Holland (1806)

Kingdom of Westphalia (1807)

Duchy of Warsaw (1807)

Kingdom of Spain (1808)

Pays allies: “allied countries whose territory was expanded by Napoleon in some cases, and continued to be governed by their native rulers”.

Most of the members of the Rhenish Confederation (including Saxony and the South-German states of Bavaria, Wurttemberg, Baden)

In four cases, this chapter deviates from the categorizations of countries made by Grab (2003). Since Grab (2003) does not include Scandinavia in his book, Finland, Sweden, and Norway are not included in his categorization. In this chapter Finland, Sweden, and Norway are regarded as *pays allies*. Portugal is categorized as *pays réunis* since Portugal was invaded by Napoleon in 1807 to ensure England could not trade on the continent (Centennia Historical Atlas).

Appendix D: Standardization (Chapter 2)

Section 2.2 shows that political pressure from Napoleon seems to have had little effect on the extent of (input and output) standardization. Therefore Section 2.3 proceeds with only differentiation. For completeness this Appendix presents the same analyses for standardization as for differentiation in the main text.

Data

Standardization on inputs

Next to the background characteristics mentioned in Section 2.3, in PISA 2006 also asked some questions to the school principal on who has authority for the textbooks, course content, teacher appointments, teacher employment conditions, and budget. These last five variables are the main independent variables, which are transformed into education system characteristics on standardization by using country averages. Table D1 shows the exact definitions of the variables and some descriptive statistics. These five standardization measures are combined into two means, namely one on the administrative standardization and one on the content related issues of standardization. These five are the variables used in Section 2.2 to look at the relation between Napoleonic pressure and education systems.

Table D1: Descriptive statistics of the standardization variables

| 2006-PISA | Mean | Standard deviation |
|---|------|--------------------|
| Responsibility for teacher hire: ratio of schools vs state | 0.11 | 0.25 |
| Responsibility for course content: ratio of schools vs state | 0.45 | 0.40 |
| Responsibility for salary increases: ratio of schools vs state | 0.41 | 0.45 |
| Responsibility for formulate budget: ratio of schools vs state | 0.79 | 0.37 |
| Responsibility for textbook use: ratio of schools vs state | 0.36 | 0.42 |
| <i>Notes:</i> Variable obtained at the school level (from the school principle). Shown mean and standard deviations are averaged country data. State means national education authority; school means teacher, principal, or school governing board. The higher the variable, the more standardization. | | |

Results

As expected, a weaker picture can be found in Table D2 which looks at the effect of standardization on student performance, than in Table 8 in Section 2.3 for differentiation. Here the first stage never has an F statistics of above 10 (the standard rule of thumb and no effect is found in the second stage for any specification for either administrative standardization or content related standardization.

The previous results on standardization often show negative effects of standardization. For instance, Hanushek, Link, and Woessmann (2011) who show, using a panel study, that decentralization of (administrative and content related) inputs has a positive effect on student performance in developed countries. Woessmann, Luedemann, Schuetz, and West (2007) find that the negative association between standardization and performance holds especially when schools are also held accountable for their outputs. Woessmann (2003) shows a positive association between content related standardization and student performance, and a mixed picture for administrative standardization.

Table D2: The effect of standardization on mathematics test scores

| | Administrative standardization | | | | | Content related standardization | | | | |
|--|--------------------------------|------------------------|----------------------|----------------------|---------------------|---------------------------------|----------------------|----------------------|----------------------|-----------------------|
| | (0) | (1) | (2) | (3) | (5) | (0) | (1) | (2) | (3) | (5) |
| 1st stage | | | | | | | | | | |
| Distance | | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| Empire | | | 0.014 (0.086) | | 0.027 (0.091) | | | 0.069 (0.047) | | 0.073 (0.051) |
| Pays réunis | | | | 0.003 (0.113) | 0.007 (0.110) | | | | 0.113 (0.067) | 0.125* (0.065) |
| Pays conquis | | | | 0.074 (0.119) | 0.0811 (0.120) | | | | -0.059 (0.071) | -0.041 (0.069) |
| Pays allies | | | | 0.052 (0.121) | 0.045 (0.127) | | | | 0.034 (0.071) | 0.015 (0.077) |
| 2nd stage | | | | | | | | | | |
| Admin | -23.230 (27.160) | -4,042.490 (76,790) | 1,326.350 (7,492) | -36.130 (145.100) | 73.580 (201.700) | | | | | |
| Content | | | | | | 12.6900 (34.4000) | 366.200 (362.400) | 378.000 (238.700) | 131.500 (92.1500) | 177.100* (100.500) |
| # of students | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 | 157,200 |
| # of countries | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 |
| R ² 2 nd stage | 0.328 | | | 0.327 | 0.280 | 0.326 | 0.043 | 0.024 | 0.294 | 0.265 |
| Partial R ² 1 st stage | | 0.00 | 0.00 | 0.02 | 0.02 | | 0.02 | 0.06 | 0.18 | 0.22 |
| F test 1 st stage | | 0.003 | 0.016 | 0.188 | 0.153 | | 1.849 | 1.908 | 2.311 | 1.912 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between standardization and student performance, instrumenting in models (2) to (5) tracking with the political pressure from Napoleon. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Model 2 is the preferred IV model. The models include imputation dummies and imputation variable interaction terms, and clustering is done on the country level. The control variables are as explained in the main text. Durbin and the Wu-Hausman tests are satisfied.

Appendix E: Full second stage model of Table 9 (Chapter 2)

| | Math | |
|--|-------------|----------|
| | Coef. | SE |
| Grade of student: 7 under modal grade | ref | |
| 6 under modal grade | -88.110*** | (14.950) |
| 5 under modal grade | 11.580* | (6.962) |
| 4 under modal grade | -81.600*** | (7.384) |
| 3 under modal grade | -187.300*** | (22.140) |
| 2 under modal grade | -133.200*** | (17.700) |
| 1 under modal grade | -73.400*** | (16.810) |
| modal grade | -23.060* | (12.960) |
| 1 above modal grade | -2.309 | (13.790) |
| 2 above modal grade | 37.980** | (18.410) |
| 3 above modal grade | -183.400*** | (16.650) |
| 4 above modal grade | -63.150*** | (17.050) |
| Gender of student: Male | ref | |
| Female | -25.400*** | (1.406) |
| Number of books in the household: 0-10 books | ref | |
| 11-25 books | 10.910*** | (1.929) |
| 26-100 books | 29.220*** | (2.963) |
| 101-200 books | 41.070*** | (3.917) |
| 201-500 books | 55.590*** | (4.190) |
| More than 500 books | 53.230*** | (4.668) |
| Track level student: General | ref | |
| Vocational | -43.080*** | (9.652) |
| School average number of books | 36.250*** | (4.261) |
| Agglomeration school: Village | ref | |
| Small town | -0.001 | (4.419) |
| Town | -3.052 | (5.157) |
| City | -5.268 | (4.812) |
| Large city | -12.640** | (5.984) |
| School size | 0.014** | (0.006) |
| Ability grouping in the school: Not for any subjects | ref | |
| For some subjects | -5.201 | (4.079) |
| For all subjects | -10.080** | (4.916) |
| GDP per capita in 2005 | -0.001 | (0.001) |
| Number of tracks | 16.330** | (7.630) |
| Constant | | √ |
| Imputation dummies and interactions | | √ |
| Observations (# of students) | 157,200 | |
| Clusters (# of countries) | 29 | |
| R ² 2 nd stage | 0.33 | |
| Partial R ² 1 st stage | 0.42 | |
| F 1 st stage | 27.61 | |
| Durbin p | 0.00 | |
| Wu-Hausman p | 0.00 | |
| <p><i>Notes:</i> The table presents coefficients (standard errors in parenthesis) on the relation between tracking and student performance, instrumenting tracking with the political pressure from Napoleon. The first stage is not depicted but can be found in Table 9 of Chapter 2. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.</p> | | |

Appendix F: The exclusion restriction (Chapter 2)

Tables F1 to F4 display the results of similar analyses as in Section 2.2 but now with the GINI coefficient, the number of police officers per 100,000 inhabitants, the expenditures on education and the proportion of women seats in parliament. These are just four examples on which Napoleon is reasonably to be expected to have had an influence. None of these is consistently found to be explained by the political pressure from Napoleon, except for the GINI coefficient. To use the political pressure from Napoleon as an instrument for education systems, Napoleon should not have influenced anything else that influences student performance today. Since Napoleon at least influenced the GINI coefficient, this exclusion restriction is violated.

Table F1: The effect of political pressure from Napoleon on the GINI coefficient

| | (1) | (2) | (3) | (4) | (5) |
|---|----------------------|----------------------|----------------------|----------------------|----------------------|
| Distance | -0.003* (0.001) | -0.003* (0.001) | -0.001 (0.001) | | -0.001 (0.001) |
| Empire | | -0.042 (1.349) | | | 0.123 (1.196) |
| Pays réunis | | | -0.959 (1.510) | -0.533 (1.325) | -0.937 (1.560) |
| Pays conquis | | | 1.363 (1.414) | 1.436 (1.389) | 1.402 (1.496) |
| Pays allies | | | -4.652** (1.737) | -5.019*** (1.610) | -4.630** (1.793) |
| Constant | 33.550*** (1.343) | 33.560*** (1.460) | 33.050*** (1.892) | 32.120*** (1.129) | 32.980*** (2.069) |
| # of countries | 26 | 26 | 26 | 26 | 26 |
| Adjusted R ² | 0.114 | 0.0752 | 0.353 | 0.371 | 0.321 |
| p F test | 0.051 | 0.156 | 0.010 | 0.004 | 0.023 |
| <i>Notes:</i> The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and GINI coefficient in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Model 2 is the preferred model. | | | | | |
| <i>Source:</i> The GINI index is the most recent number (2008-1999) from the World Bank. | | | | | |

Table F2: The effect of political pressure from Napoleon on the number of police officers per 100,000 inhabitants

| | (1) | (2) | (3) | (4) | (5) |
|--|------------------------|------------------------|------------------------|------------------------|----------------------|
| Distance | -0.050 (0.033) | -0.052 (0.034) | -0.007 (0.041) | | -0.006 (0.041) |
| Empire | | 27.040 (33.470) | | | 30.540 (33.560) |
| Pays réunis | | | 8.238 (40.450) | 10.950 (36.640) | 14.920 (41.260) |
| Pays conquis | | | -9.369 (40.790) | -8.641 (39.780) | -0.979 (41.960) |
| Pays allies | | | -97.130* (49.340) | -100.000** (45.560) | -94.330* (49.620) |
| Constant | 364.500*** (35.370) | 353.700*** (38.030) | 343.300*** (51.450) | 336.000** * | 323.000** * |
| # of countries | 29 | 29 | 29 | 29 | 29 |
| Adjusted R ² | 0.043 | 0.03 | 0.09 | 0.13 | 0.09 |
| p F test | 0.144 | 0.254 | 0.178 | 0.094 | 0.218 |
| <p><i>Notes:</i> The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and the number of police officers per 100,000 inhabitants in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Model 2 is the preferred model.</p> <p><i>Source:</i> The number of police officers (per 100 000 inhabitants) for 2008 is from Eurostat.</p> | | | | | |

Table F3: The effect of political pressure from Napoleon on the expenditure on education

| | (1) | (2) | (3) | (4) | (5) |
|-------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Distance | 0.0001** (0.000) | 0.001** (0.000) | 0.001* (0.000) | | 0.001* (0.001) |
| Empire | | -0.268 (0.378) | | | -0.295 (0.411) |
| Pays réunis | | | 0.185 (0.492) | -0.155 (0.475) | 0.120 (0.505) |
| Pays conquis | | | -0.176 (0.496) | -0.267 (0.515) | -0.257 (0.514) |
| Pays allies | | | 0.0234 (0.600) | 0.387 (0.590) | -0.004 (0.608) |
| Constant | 4.359*** (0.398) | 4.466*** (0.429) | 4.272*** (0.626) | 5.186*** (0.383) | 4.467*** (0.689) |
| # of countries | 29 | 29 | 29 | 29 | 29 |
| Adjusted R ² | 0.12 | 0.11 | 0.03 | -0.06 | 0.01 |
| p F test | 0.034 | 0.087 | 0.331 | 0.693 | 0.411 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and the expenditures on education in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Model 2 is the preferred model.

Source: The expenditures on education are from the World Bank (Public spending on education, total (% of GDP)) for 2005.

Table F4: The effect of political pressure from Napoleon on the number of women seats in parliament

| | (1) | (2) | (3) | (4) | (5) |
|-------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Distance | 0.003 (0.004) | 0.003 (0.004) | -0.001 (0.004) | | -0.001 (0.004) |
| Empire | | -0.813 (3.711) | | | -0.249 (3.530) |
| Pays réunis | | | 4.502 (4.181) | 4.702 (3.786) | 4.448 (4.340) |
| Pays conquis | | | 2.759 (4.216) | 2.813 (4.110) | 2.691 (4.414) |
| Pays allies | | | 14.280*** (5.100) | 14.060*** (4.708) | 14.260** (5.219) |
| Constant | 24.570*** (3.877) | 24.900*** (4.217) | 22.420*** (5.317) | 21.890*** (3.057) | 22.590*** (5.915) |
| # of countries | 29 | 29 | 29 | 29 | 29 |
| Adjusted R ² | -0.01 | -0.05 | 0.15 | 0.18 | 0.11 |
| p F test | 0.440 | 0.729 | 0.101 | 0.048 | 0.179 |

Notes: The table presents coefficients (standard errors in parenthesis) on the relation between political pressure from Napoleon and the number of women seats in parliament in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Model 2 is the preferred model.

Source: The proportion of seats held by women in national parliaments (%) for 2008 is from the World Bank.

Appendix G: Analyses using data from 1970 (Chapter 2)

This appendix shows analyses similar to those done in Section 2.4, but now using data on tracking and student performance from 1970 instead of 2006.

Data

Student and school level data

Student performance and student and school level controls are obtained in 1970 from the International Association for the Evaluation of Educational Achievement (IEA) First International Science Study (FISS). FISS is the second internationally comparable study, performed by the IEA.¹⁸ For 2006, as before, I use data from the Program for International Student Assessment (PISA).

For FISS 1970 data was collected in 19 countries and regions, of which ten are used in this paper. These ten countries, or regions, are Flemish Belgium, French Belgium, England, Finland, West-Germany, Hungary, Italy, the Netherlands, Scotland, and Sweden. Together this amounts to 25,409 students in 1033 schools. The more extended PISA 2006 sample used

18. In 1964 another study was held by the IEA but no information on standardization was collected and only 6 European countries participated.

in the main text of Chapter 2 contains the ten countries used in FISS. In these ten countries there are 41,790 students in 1,671 schools who participated in PISA 2006. The age target of both data sets is quite similar: FISS 1970 were 14-year-old students, while PISA 2006 were 15-year-old students. Like in the main text, missing values are replaced by either school or country averages and the analyses are done with student weights and clustering of errors on country level. This clustering on country level implies that in the restricted sample only 10 observations exists.

Both FISS and PISA contain a science test score, which is in this paper standardized over all countries to be able to compare the results over time. So while in the main text, I focus on the mathematics test score, here I focus on science since FISS only contains a science test score. FISS and PISA also provide, on a limited scale, similar information on the student and the school. Both data sets contain data on the controls used in the main text: the gender of the student, the number of books in the household (as a proxy for social economic background), the grade of the student, the school size, the location of the school, the track type of the school or students, and whether the school has ability grouping.

Tracking

For tracking in 1970 I constructed one differentiation measure based upon national reports on the state of education from UNESCO in the early seventies (or late sixties). For differentiation in 1970 a combination of the number of tracks and age of selection is made. Table G1 shows differentiation in 2006 and 1970 for the sample of ten countries.

Some countries made changes in the level of differentiation in their education systems between 1970 and 2006 (see also Chapter 5). For instance, political opinion in England, Sweden and Finland in 1970 was in favor of a comprehensive system and thus in the years following 1970 these countries saw a move towards a comprehensive system. However, in 1970 Scotland was the only country which had already a comprehensive system in place for its 14-years-pupils. England, Sweden and Finland followed in the years after 1970. Another example of a country that changes its level of differentiation is Belgium which had a less differentiated system in 1970, with only a general and vocational track and selection at age 14, than in 2006, with four tracks and selection at age 12.

Table G1: Differentiation in the restricted sample

| Country | Number of tracks in 2006 | Age of first selection in 2006 | Differentiation in 2006 | Differentiation in 1970 |
|-----------------|--------------------------|--------------------------------|-------------------------|-------------------------|
| Scotland | 1 | 16 | 0 | 0 |
| England | 1 | 16 | 0 | 1 |
| Sweden | 1 | 16 | 0 | 1 |
| Finland | 1 | 16 | 0 | 1 |
| Italy | 3 | 14 | 2 | 3 |
| Hungary | 3 | 11 | 3 | 3 |
| Flemish Belgium | 4 | 14 | 2 | 2 |
| The Netherlands | 4 | 12 | 3 | 3 |
| (West-)Germany | 4 | 10 | 3 | 3 |
| French Belgium | 5 | 14 | 3 | 2 |

Source: Differentiation in 1970 is from the 1968 yearbook of education of UNESCO and the different national reports of UNESCO of 1968, 1971, 1973, and 1975. Differentiation in 2006 is a combination of the number of tracks in 2006 and the age of first selection in 2006. See Table 1 of Chapter 2 for further sources.

Results

Table G2 presents the results on differentiation using the FISS 1970 restricted sample and a comparable dataset of PISA 2006, containing the same 10 countries¹⁹. The full 2006 sample revealed a significant positive effect of the number of tracks on student performance and an insignificant negative effect of the age of first selection. The restricted 2006 sample shows a weaker relation between both the age of first selection and the number of tracks and student performance: The coefficient on the number of tracks is still positive, although insignificantly, but the coefficient on the age of first selection is no longer consistently negative (not shown, but available on request). Table G2 shows that using this restricted sample in both 1970 and 2006 countries that were conquered by France have more differentiation. And for the combined measure of differentiation there is no effect on student performance in 2006 (right panel) although the first stage is still strong, while there is a consistent negative effect in 1970 (left panel). The first stage shows similar results as before, with only a changing sign for *pays conquis*.

In 1970 one step up the differentiation measure leads to one-tenth to one quarter standard deviation lower science test score. Similarly to before, this is equivalent to a move from a household with no books to a household with 1-10 books (.265***).²⁰ Again, it should be kept in mind, that, due to the clustering, these models effectively only have ten observations.

19. West-Germany is Germany minus the states Berlin, Saxony-Anhalt, Brandenburg, Mecklenburg-Vorpommern, Saxony and Thuringia.

20. In 1970 the variable number of books in the household has five categories: none, 1-10 books, 11-25 books, 26-50 books, and 51 or more.

Table G2: The effect of tracking on science test scores using the FISS 1970 and restricted PISA 2006 sample

| Model: | 1970 | | | | | 2006 | | | | |
|--|--------------------|----------------------|---------------------|---------------------|------------------------|------------------|----------------------|----------------------|---------------------|---------------------|
| | (0) | (1) | (2) | (3) | (4) | (0) | (1) | (2) | (3) | (4) |
| 1st stage | | | | | | | | | | |
| Distance | | -0.001*** (0.000) | -0.001** (0.000) | 0.000 (0.000) | -0.025*** (0.005) | | -0.001*** (0.000) | -0.001*** (0.000) | -0.001** (0.000) | -0.001** (0.000) |
| Empire | | | 0.632 (0.425) | | 54.840*** (10.440) | | | 1.029 (0.659) | | -0.955 (0.542) |
| Pays reunis | | | | 0.898* (0.422) | -84.270*** (16.010) | | | | 1.218* (0.546) | 2.441*** (0.689) |
| Pays conquis | | | | 1.117*** (0.177) | 33.150*** (6.060) | | | | 0.386 (0.505) | -0.195 (0.454) |
| Pays allies | | | | -0.048 (0.174) | -19.630*** (3.739) | | | | 0.833 (0.455) | 1.092** (0.393) |
| 2nd stage | | | | | | | | | | |
| Differentiation | -0.083* (0.045) | -0.231 (0.147) | -0.251** (0.114) | -0.101* (0.060) | -0.117** (0.050) | 0.010 (0.062) | 0.056 (0.136) | 0.078 (0.116) | 0.006 (0.071) | -0.009 (0.065) |
| # of students | 25,409 | 25,409 | 25,409 | 25,409 | 25,409 | 73,171 | 73,171 | 73,171 | 73,171 | 73,171 |
| # of countries | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| R ² 2 nd stage | 0.230 | 0.215 | 0.211 | 0.229 | 0.229 | 0.353 | 0.351 | 0.348 | 0.353 | 0.353 |
| Partial R ² 1 st stage | | 0.214 | 0.279 | 0.704 | 0.874 | | 0.438 | 0.547 | 0.881 | 0.907 |
| F test 1 st stage | | 13.74 | 5.219 | 43.12 | 46.45 | | 40.75 | 11.22 | 126.6 | 108.1 |

Notes: Control variables are a constant, the gender of the student, the number of books in the household, the grade of the student compared to the modal grade in the country, the school size, the location of the school, the track type of the school or students, and whether the school has ability grouping. Durbin and the Wu-Hausman tests are satisfied, except for model 5 in 2006. Clustering on country level. Imputation dummies and interactions are included. Model 2 is the preferred model.

Appendix H: Nonlinear models (Chapter 3)

This Appendix presents the same models as Table 4 and 6 in the main text of Chapter 3 but now including the number of tracks in a country as dummy variables instead of including it as a continuous variable. To avoid too many dummies and, especially, too many interactions between the number of tracks and whether school principals consider prior performance when accepting students to the school, we split up the categorical variable in three: No tracking, two or three tracks, or four or five tracks. The results on the relation between tracking, selection and student performance are qualitatively the same as in the main text, as can be seen in Table H1: The interactions between number of tracks and whether school principals consider prior performance sometimes or always are positive and significant. The combined coefficients of the main effects and the interaction effects show that for students in schools where the principal considers prior performance in a country with four or five tracks, tracking is either positive or neutral as compared to students in a school where the principal does not consider prior performance in a country without tracking. Different to the main results is that *students in a country with two or three tracks always perform much worse than students in any of the education systems*, irrespective whether their school principal considers prior performance. This is further examined in Section 3.6.4.

Table H1: Non-linear models looking at student performance

| Dependent variable | Reading | Mathematics | Science |
|--|---------------------|----------------------|----------------------|
| School considers prior performance | | | |
| <i>Sometimes</i> | -6.56 (2.53) | -4.69** (2.36) | -5.83*** (2.15) |
| <i>Always</i> | -4.97 (3.21) | -4.04 (3.77) | -5.47 (4.93) |
| 2-3 tracks | -32.61*** (8.87) | -39.10*** (10.38) | -36.67*** (11.77) |
| 4-5 tracks | -18.08** (8.32) | -2.98 (10.23) | -10.45 (9.33) |
| Sometimes*2-3 tracks | 6.11 (3.99) | 3.82 (4.11) | 6.00 (4.23) |
| Always*2-3 tracks | 9.24 (6.83) | 3.39 (6.07) | 9.58 (8.22) |
| Sometimes*4-5 tracks | 10.58*** (3.95) | 8.78** (4.18) | 2.45*** (4.11) |
| Always*4-5 tracks | 21.79*** (5.48) | 21.92*** (5.14) | 26.74*** (6.28) |
| (pseudo)LL | -43755 | -43732 | -44088 |
| # of students | 187768 | 187,768 | 187,768 |
| # of countries | 31 | 31 | 31 |
| <i>Notes:</i> Coefficients with standard errors in parenthesis. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See also notes of Table 3 of Chapter 3. | | | |

Table H2 shows similar models to those in the first 3 columns of Table 6 but also here including the number of tracks in a country as three dummy variables. The interactions between parental background and the number of tracks are significantly negative. This means

that the relation between parental background and student performance is lower in countries with more tracks. Also this result is similar to those presented in the main text.

Table H2: Non-linear models looking at equality of opportunity.

| Dependent variable | Reading | Mathematics | Science |
|--|---------------------|----------------------|----------------------|
| School considers prior performance | | | |
| <i>Sometimes</i> | -1.68 (2.02) | -1.08 (1.88) | -0.62 (2.08) |
| <i>Always</i> | 6.87* (3.28) | 5.76 (2.98) | 8.24* (3.67) |
| 2-3 tracks | -30.51*** (8.52) | -39.43*** (10.42) | -35.00*** (10.99) |
| 4-5 tracks | -6.95 (7.73) | 8.40 (10.08) | 3.12 (9.42) |
| Parental background* 2-3 tracks | -12.50*** (2.34) | -12.66*** (2.56) | -14.18*** (2.55) |
| Parental background* 4-5 tracks | -10.90*** (3.29) | -10.59*** (3.30) | -11.33*** (3.40) |
| (pseudo-)LL | -43742 | -43719 | -44074 |
| # of students | 187768 | 187,768 | 187,768 |
| # of countries | 31 | 31 | 31 |
| <i>Notes:</i> Coefficients with standard errors in parenthesis. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See also notes of Table 3 of Chapter 3. | | | |

Appendix I: Comparing different models for reading and science (Chapter 3)

This appendix compares the different estimation models for reading and science, like is done in the main text for mathematics in Table 5 of Chapter 3. The main difference is that the models showing results using only the between country variation is not significant for reading or science. However they do show the same trend and it must be remembered that since these models use country level data only the number of observations is effectively reduced to 31. An F test is performed to see whether the main and interaction effects are jointly significant and they are for mathematics (p-value= 0.02) and reading (p=0.00). For science the five variables are not jointly significant, but the three variables relating to schools that always consider prior performance are jointly significant at the five percent level (p-value= 0.04).

Table I1: Comparing different models on student performance (reading as dependent variable)

| | (1) Between countries, between schools | (2) Within country | (3) Between countries |
|--|--|-----------------------|--------------------------|
| School considers prior performance | | | |
| <i>Sometimes</i> | -5.97** (2.54) | -4.37** (2.17) | 35.24 (39.05) |
| <i>Always</i> | -10.13** (4.11) | -2.67 (4.10) | -66.67** (30.42) |
| Number of tracks (0-4) | -5.74** (2.70) | - | 5.47 (12.54) |
| Sometimes*Number of tracks | 2.83** (1.15) | 2.54** (1.05) | -36.50 (23.61) |
| Always*Number of tracks | 8.64*** (1.82) | 6.03*** (1.96) | 19.95 (18.06) |
| Country FE | - | √ | - |
| (pseudo)LL | -43759 | -43727 | -43758 |
| # of students | 187768 | 187768 | 187768 |
| # of countries | 31 | 31 | 31 |
| <i>Notes:</i> Coefficients with standard errors in parenthesis. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Column (1) shows the main model as depicted in column (2) of Table 4. Column (3) measures the school variables “school principal consider prior performance” on a national level and thus depicts the proportion of schools (between 0 and 1) in the country with school which say they always or sometimes consider prior performance. See also notes of Table 3 of Chapter 3. | | | |

Table I2: Comparing different models on student performance (science as dependent variable)

| | (1) Between countries, between schools | (2) Within country | (3) Between countries |
|---|--|-----------------------|--------------------------|
| School considers prior performance | | | |
| <i>Sometimes</i> | -5.99** (2.44) | -5.40** (2.20) | 78.61 (55.61) |
| <i>Always</i> | -10.90** (5.06) | -5.07 (4.88) | -86.87* (46.45) |
| Number of tracks (0-4) | -2.89 (3.31) | - | 6.40 (16.08) |
| Sometimes*Number of tracks | 3.45*** (1.08) | 3.33*** (0.99) | -42.89 (32.60) |
| Always*Number of tracks | 9.68*** (2.04) | 7.39*** (2.08) | 34.62 (24.08) |
| Country FE | - | √ | - |
| (pseudo)LL | -44092 | -44054 | -44092 |
| # of students | 187768 | 187,768 | 187,768 |
| # of countries | 31 | 31 | 31 |
| <p><i>Notes:</i> Coefficients with standard errors in parenthesis. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Column (1) shows the main model as depicted in column (2) of Table 4. Column (3) measures the school variables “school principal consider prior performance” on a national level and thus depicts the proportion of schools (between 0 and 1) in the country with school which say they always or sometimes consider prior performance. See also notes of Table 3 of Chapter 3.</p> | | | |

Appendix J: Subsample of early tracking countries (Chapter 3)

Table J1 presents similar results to Table 4 of Chapter 3 for late tracking countries only, since in these countries the time in a tracked system may have been too short to have an effect on student performance. The results are for the largest part comparable to those using the full sample, but the significance of the results is lower, especially for mathematics and science.

Table J1: Results comparable to Table 4 of Chapter 3 for late tracking countries only

| | (1) Reading | (2) Math | (3) Science |
|--|--------------------|--------------------|--------------------|
| School considers entrance requirements | | | |
| <i>Sometimes</i> | -6.86*** (2.20) | -4.51*** (2.16) | -5.37*** (2.01) |
| <i>Always</i> | -7.61 (4.78) | -5.84 (5.25) | -6.34 (6.34) |
| Number of tracks (0-4) | -7.85** (3.71) | -7.94** (3.96) | -5.15 (4.64) |
| Sometimes*Number of tracks | 3.32** (1.62) | 2.78 (1.90) | 2.17 (1.77) |
| Always*Number of tracks | 6.31** (2.65) | 3.63 (2.25) | 5.01* (2.76) |
| (Pseudo) LL | -28,138 | -28,132 | -28,358 |
| # of students | 131,423 | 131,423 | 131,423 |
| # of countries | 14 | 14 | 14 |
| <i>Notes:</i> The table presents coefficients from random effects models (standard errors in parenthesis) on the relation between student performance and whether or not schools consider prior performance when selecting students and the number of tracks in a country. The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The control variables are as described in the text. The models include imputation dummies and imputation variable interaction terms. | | | |

