

The education revolution on horseback II : using the Napoleonic wars to elicit the effect of tracking on student performance

Citation for published version (APA):

Korthals, R. A. (2016). *The education revolution on horseback II : using the Napoleonic wars to elicit the effect of tracking on student performance*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda No. 015 <https://doi.org/10.26481/umagsb.2016015>

Document status and date:

Published: 01/01/2016

DOI:

[10.26481/umagsb.2016015](https://doi.org/10.26481/umagsb.2016015)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Roxanne Korthals

**The Education Revolution on
Horseback II: Using the
Napoleonic Wars to Elicit the
Effect of Tracking on Student
Performance**

RM/16/015

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O. Box 616
NL- 6200 MD Maastricht
The Netherlands

The Education Revolution on Horseback II: Using the Napoleonic Wars to Elicit the Effect of Tracking on Student Performance¹

Roxanne Korthals

Abstract: Previous literature has found inconsistent effects of tracking students in secondary school on student performance using various ways to alleviate the endogeneity in tracking. 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. I use this theory to estimate the effect of tracking on student performance in Europe, instrumenting tracking by the political pressure caused by the Napoleonic Wars. The relation between political pressure by Napoleon and tracking is strong and leads in the second stage to a consistent positive effect of tracking on student performance. One important limitation of this analysis is that it is reasonable to assume that political pressure from Napoleon influenced many facets of European countries.

JEL Codes: I2

Keywords: tracking, educational policy, PISA

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 and the IPN Kiel for providing me with the extended PISA data for Germany. I further would like to thank Francisco Ramirez, David Figlio, Bas ter Weel, Ulf Zoelitz, Andreas Ammermüller, 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. This paper is the second of two companion papers on Napoleon and tracking students in secondary education. The first is Korthals (2016).

1. Introduction

One aspect of the design of secondary education is whether or not students are formally separated into distinct educational programs (so-called tracks) based on ability level. With tracking, students of different tracks have their own curriculum and the students are often not in the same building, or even the same school. It is often believed that tracking students into distinct educational programs leads to lower mean student performance and to less intergenerational mobility (OECD, 2010). However, studying the effect of tracking is complicated since whether or not students are tracked in secondary school is determined on the country (or state) level, and tracking itself might not be the only important country level factor influencing student performance. Any effect attributed to tracking in naïve estimations might therefore be partly caused by unknown third factors.

The aim of this paper is to use an instrumental variable technique to estimate the effect of tracking on student performance. The instrument employed in this paper is based on Ramirez and Boli (1987). Ramirez and Boli (1987) propose a theory that points to an explanation for the design of 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. The political pressure from France, as theorized of Ramirez and Boli (1987), is related to the level of tracking, both seen by the number of tracks available to fifteen-year-old students and by the age of first selection. The pressure from France is therefore used as an instrument to alleviate the endogeneity between tracking and student performance. Naturally, the Napoleonic Wars in the 1800s influenced many aspects of European societies and thus these (causal) analyses should be interpreted with caution. Given this consideration, the first stage of the analyses, looking at the relation between the political pressure from Napoleon, is strong. In the second stage, positive effects are found for the number of tracks in a country on student performance. Negative, but insignificant, effects are found for the age of first selection.

The existing literature trying to alleviate the endogeneity of tracking is not extensive. Hanushek and Woessmann (2006) provide international difference-in-difference analyses into the effects of selection into tracks at an early age and find indications of a negative effect of tracking at an early age on country averaged student performance. This result is confirmed by Van Elk, van der Steeg and Webbink (2011) who use regional variation in the age of tracking in the Netherlands. Ariga and Brunello (2007) instrument the tracking length and

find a positive effect of the time spend in tracks. Kerr, Pekkarinen, and Uusitalo (2013) use a policy reform in Finland and find insignificant effects for logical reasoning and arithmetic test scores and a small positive effect on verbal test scores. This paper adds to this literature by using a new instrument to alleviate the endogeneity of tracking.

This paper is structured as follows: the next section elaborates on the theory by Ramirez and Boli (1987) and the used instrument. Section 3 and 4 describe the estimation method and the data. Section 5 then shows the results of the instrumental variable method to estimate the effect of tracking on student performance. Section 6 pays more attention to the exclusion restriction, while Section 7 runs a number of robustness checks. Section 8 concludes.

2. Napoleon and Ramirez and Boli (1987)²

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. The companion paper to this paper (Korthals, 2016) provides more historical background on the rise of Napoleon as emperor of France and on the wars on the European continent during his reign.

² This section is based on Korthals (2016), the companion paper of this paper.

In their 1987 paper *The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization*. Ramirez and Boli present seven case studies with which the authors show that the then real threat of war with and annexation by the French in the 1800s induced European countries to create a national identity among their population. This national identity was important since it increased the public's (physical and mental) willingness to pay and fight for the war with France. The introduction of mass public educational systems by the national state helped the state in forming this national identity since school curricula offer the opportunity to educate the people in the cultural and political tradition of the country. Consequently, it was the threat of war with the French army that induced European countries to set up mass public education systems.

The theory of Ramirez and Boli (1987) can best be described by a quote from their paper: "Our view is that European states became engaged in authorizing, funding, and managing mass schooling as a part of the endeavour 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. 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. The South of Italy acted like France and England, while the North made, like Prussia and Denmark, earlier steps towards a mass public education system. Regions that were under external pressure (by Napoleon in most cases) were induced to introduce mass public education systems while "the most dominant powers [ed: France and England] were able to resist the system wide pressures favouring mass education [...]" (Ramirez and Boli, 1987, p. 4). More on these case studies can be found in the companion paper of this paper, Korthals (2016).

3. Estimation Method

The IV models estimated in this paper are 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 ($TRACKING_c$) related to this pressure. In the second stage, I then use the fitted values ($\widehat{TRACKING}_c$) to estimate the effect of tracking on student performance ($TEST SCORE_{i,c}$). I

control for individual, school and country characteristics captured in $\mathbf{X}_{i,c}$. 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, and GDP per capita in 2005. The variable of interest is β which displays the effect of tracking on student performance.

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

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

For a valid IV two conditions must hold: First, the instrument should be related to the endogenous variable. Second, the instrument should not be related to anything else, observable or not, that might influence the outcome, in this case, student performance. The second condition (the exclusion restriction) cannot be tested directly and must be taken on good faith. Section 6 and Appendix B elaborate more on this. The first condition can be verified by a strong first stage. Section 5 shows that this first condition is indeed fulfilled.

4. Data

4.1 Individual level data on student performance

Student performance data is obtained from the Programme for International Student Assessment (PISA) 2006 organized by the OECD. PISA 2006 was conducted in 57 countries, of which 21 are in Europe and located on the battle ground of the Napoleonic Wars. France is excluded 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 Naples³), and Germany (into Bavaria, Hanover, Prussia, Rhineland, Saxony, Westphalia).⁴ This paper therefore uses 29 regions, which will be called countries for simplicity, 6578 schools, and 157,200 students.⁵ See Table 1 for a full list of countries and Appendix A for a more extensive description of the regions. Analyses are done with student weights and clustering of errors on country level. Using country level variables on tracking

³ Some Italian regions are excluded since in PISA they are grouped together in a non-geographical way.

⁴ See Appendix A for a more extensive description of the regions.

⁵ Students in schools with less than 5 students in PISA are excluded, as are migrant students, or students who have missing values on their age, gender, or socio-economic status. Other missing values are replaced by either school or country averages and imputation dummies and imputation dummy interactions are included in all models. The amount of missing variables ranges from 0.000% (student in vocational education) to 0.042% (school size).

and clustering on the country level implies that the sample effectively only contains 29 independent observations even though individual level data is used.

PISA 2006 contains a reading, mathematics, and science test score. This paper 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 (see also Section 7 Robustness). Next to the test scores, PISA also provides information on the student and the school of which the following are used as control variables: the gender of the student, the number of books in the household (as a proxy for parental background), the grade of the student, the track type of the student, the school average of the number of books, the school size, and the location of the school. In Section 7 (Robustness) also models without controls are shown.

4.2 Country level data on tracking

Two aspects of tracking in education systems are used: the number of tracks available to 15-year-old pupils and the age of first selection. These measures are obtained from OECD (2007, table 5.2) and adjusted for the German states using Woessmann (2007) and Kultusminister Konferenz (2013) and for French and Flemish Belgium using Eurydice (2013). Table 1 and Figure 1 show the 2006 levels of tracking for the countries in the sample. 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 shows the level of tracking of current day European countries. The geographical spread in Figure 1 suggests a weak relation between France and education systems. Countries in the periphery of Europe have less tracking than countries in the inner region.

--- Table 1 here ---

--- Figure 1 here ---

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

paper and OECD (2010) is unravelled in Table 2. First, the OECD correlation is based on PISA 2000, while in this paper 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.⁶ This results in a correlation between the age of first selection and student performance of -0.23. The OECD correlation seems to be mainly driven by Mexico and the use of the collapsed variable age of first selection, and can thus be considered a statistical artefact.

--- Table 2 here ---

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”, or in other words they are very highly correlated. This justifies the use of a single theory for the instruments for both endogenous variables. 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 be to use more instruments for both aspects simultaneously. In the discussion of the results, this is further addressed.

4.3 Country level data on the political threat from Napoleon

⁶ 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 paper are two differences: Switzerland selects at age 14 (instead of age 12) and Poland selects at age 16 instead of 15.

Political pressure from Napoleon Bonaparte and the French army is captured by two *ex-ante* measures and one *ex-post* measure. The *ex-ante* political threat by Napoleon captures the theoretical threat, while the *ex-post* measure relates to which countries were actually annexed and not which countries were under threat. The *ex-post* measure therefore misses out on the countries that were under threat but were not defeated or conquered. The first, *ex-ante*, way to capture the political threat from Napoleon is by the distance of one's own capital to the French capital Paris. This gives an indication of how easy it was to move troops to the battlefield to execute the threat. 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. Brussels, the capital of Belgium, is closest to Paris with a distance of 264 km, while Turku, the capital of Finland, is furthest away with 2304 km.

The second, *ex-ante*, measure is whether or not countries belonged to one of the following empire: Russian Empire, the Habsburg Empire, and Prussia.⁷ 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 or it would have increased the threat when there were struggles between Napoleon and the empire in question. To capture this influence from belonging to an empire, I include a dummy for countries that belonged to the Russian Empire, the Habsburg Empire, or Prussia.⁸ The empires are shown in Figure 2.

--- Figure 2 here ---

Finally, a categorization of European countries by Grab (2003) is used as a measure of the *ex-post* threat. For instance, 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*

⁷Prussia was a kingdom and not an empire but it is referred to one in this paper 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 paper.

⁸ Prussia was a Kingdom and not an empire but it will be referred to one in this paper 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 paper.

conquis), like Spain and Switzerland. However, not all territories in Europe were at one point under French control. Figure 3 shows how Europe was divided under Napoleons reign.

--- Figure 3 here ---

5. Results

Table 3 shows the first and second stages for both the number of tracks and the age of first selection. The upper panel of Table 3 shows the first stage, while the lower panel shows the second stage. The table shows different first stage models using the distance to Paris, whether countries belonged to an empire, and the categorization by Grab (2003) to instrument tracking in a country. The first stage of IV model 1 shows the relation between distance and the number of tracks available (age of first selection) and shows a strong negative (positive) correlation, meaning that the further away one's own capital is from Paris the fewer tracks (the later selection) its education systems will have. IV 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-three percent of the variation in number of tracks and twenty-three percent of the variation in the age of first selection. IV model 2 shows that countries belonging to an empire in 1800 have more tracks available to students than countries not belonging to an empire in 1800. IV model 3 also includes the Grab dummy and it shows that *pays réunis* have more tracks and that *pays allies* select their students earlier. The first stage models do remarkably well: All have an F statistic of above 10 and they explain twenty-one to fifty percent of the variation. The estimates say that countries that have their capital further away from Paris have fewer tracks available to pupils nowadays and select their students later and that countries that were part of an empire have more tracks.

--- Table 3 here ---

The negative correlation between the age of first selection and student performance shown in Table 2 is most probably affected by endogeneity. This paper tries to alleviate this endogeneity by instrumenting tracking by the political pressure from Napoleon. The lower panel of Table 3 shows the OLS and IV results of tracking on student performance. The first dimension of tracking, the number of tracks, has a significant effect on student performance as seen by the mathematics test scores. The OLS model reveals that one more track available

to students leads to a 9.0 point (**) increase in student math test scores. IV models 2 and 3 show that this estimate is downwards biased since the IV models reveal a coefficient of 16.45**-17.00** of the number of tracks. 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 IV model 2 says that when a system has one more track, student performance as measured by a math test score will be about one-sixth of a standard deviation higher (mean: 507; sd: 93.1). This seems small, but it is 1.5 times the difference in the relation between 0 to 10 books in the household and 11 to 25 books (11.06***) which is quite substantial.⁹

The second dimension of tracking, the age of first selection, has no significant effect on student performance, as shown by the right lower panel of Table 3, although all estimates are consistently negative and quite substantial. The lack of significance is not unexpected since the first stage shows a weak relation between the age of first selection and the instrument. On the other hand, 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.¹⁰ 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.

6. The 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

⁹ 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.

¹⁰ IV models using more than two instruments for both endogenous variables simultaneously resulted into similar coefficients (not shown). However, the standard errors were much larger leading to insignificant results. This indicates that with only 29 countries this simultaneous estimation is too demanding for the data.

correlated events could potentially bias the IV results. The first stage analyses 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.¹¹ Table 4 shows the relation between the political pressure from Napoleon and these outcomes, using IV model 2 of Table 3. What can be seen is that in general political pressure from Napoleon is not really related to these current day societal factors: The individual instruments have in two cases a significant effect, but in most instances they do not. However, the F test on joint significance shows the model on educational spending is significant (at the 10 percent level).

--- Table 4 here ---

Appendix B also shows the other first stage models from Table 3 using these societal characteristics as outcomes. Both the expenditures on education and the GINI coefficient have a significant relation with the political pressure from Napoleon in two out of three IV models. This suggests that at the very least pressure from Napoleon influenced (determinants of) the GINI coefficient and the expenditures on education which might have influenced educational opportunities and other aspects of education which have an effect on student performance and thus the exclusion restriction is violated.

7. Robustness

7.1 Other test subjects

The main results presented in this paper are for the mathematics test scores. Table 5 shows (a shortened version of) the same analyses using the reading and the science test scores, with the mathematics results displayed as a benchmark. The results using science test scores are very similar to those using the mathematics test scores. For reading, the results are less significant and also the coefficients are smaller.

--- Table 5 here ---

7.2 Sample selection

¹¹ More information on this is given in Appendix B.

Another issue could be the sample used in this paper. For instance, I split up Belgium, Italy, Germany, and the UK into regions, which increases my number of observations but might also lead to counting observations double if these regions are not different from each other as I assume. In Table 6 results are presented in which only full countries are used. Per model the estimated coefficients are shown with significance levels. The F statistics is shown below each coefficient in parenthesis. However, the main results using the 29 regions (row 1) are very similar to the results in which the regions are collapsed into the 20 countries (row 2).

--- Table 6 here ---

A number of countries in Europe were excluded all together from my sample because they do not lie in the sphere of influence of Napoleon and therefore could not be (positively or negatively) affected by the political pressure from Napoleon. Row 3 in Table 6 shows results with an extended sample in which I also include Bulgaria, Estonia, Greece, Ireland, Iceland, Lithuania, Latvia, Montenegro, Romania, Russia, Serbia, and Turkey. The results are even stronger using the extended sample.

It could also be that one specific country drives the results. I therefore excluded each country one by one. In most cases this hardly made any difference. The coefficients were very similar and no country made the F statistics drop below 10 for the analyses using number of tracks. Table 7 shows the most extreme deviations from the main results on the number of tracks. Excluding Flanders Belgium leads to the lowest coefficients, while excluding Finland leads to the largest increase in the coefficients.

--- Table 7 here ---

When I do the same for the age of first selection, the results only truly change when Finland is excluded. When Finland is excluded the coefficients are 1.5 to 1.75 times the size of before and now all three models are significant (model IV1: -7.921 vs -14.91**; model IV2: -10.22 vs -16.12**; model IV3 -6.427 vs -9.855**). However, the F statistics is small and even under 10 for model IV2 and model IV3.

7.3 Controls

In the main analyses I control for student, school, and country characteristics. Table 8 presents results without controls or with different controls. First, I add controls for the education expenditures (row 2), or the GINI coefficient (row 3), or both (row 4) since those two variables are revealed in Section 6 (The exclusion restriction) to also have a relation with political pressure from Napoleon. The results are robust to the inclusion of these variables, which is reassuring. The inclusion of more country level variables lower the F statistic of the first stage, but it mostly stays above the required level of 10.

Second, I exclude all controls (row 5). The coefficients are much smaller, while the standard errors remain similar in size which renders the coefficients not significant. This is stronger for the number of tracks as outcome, than the age of first selection.

Third, I include only student controls or only country controls. Including the student controls leads to very similar results as before. Using only GDP per capita as a control leads to similar results as without any controls. Controlling also for the GINI coefficient and the education expenditures leads again to very similar results as the main results.

--- Table 8 here ---

7.4 Nonlinear effects

In the main model the number of tracks and the age of first selection are included as continuous variables. It could be that the results are different for certain non-linear thresholds. Table 9 shows results where the endogenous variable is a dummy for countries which have more than 1 track (row 1), more than 2 tracks (row 2), or where selection takes place before age 12 (row 3), or before age 14 (row 4). The coefficients are naturally much larger since now the endogenous variable is a dummy and not a variable with multiple values. Table 9 shows that for the number of tracks especially more than 2 tracks has a positive effect on mathematics scores. The F statistics for the non-linear age of first selection is very low.

--- Table 9 here ---

8. Conclusion

This paper finds a positive effect of tracking students in secondary school on student performance, as measured with the mathematics test scores of PISA 2006. It employs an

instrumental variable approach by using the political pressure from Napoleon Bonaparte in the 19th century as exogenous variation that is related to the level of tracking in education systems. The use of the political pressure from Napoleon as an instrument for the level of tracking is based upon the 1987 paper of Ramirez and Boli: *The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization*. Ramirez and Boli (1987) state that external challenges, of which Napoleon was one of the most pronounced, caused European states to introduce mass public education systems. Mass public education systems were set up to ensure all people were united as citizens of the country. The idea was that this feeling of citizenship would make the public willing to fight and pay for the ongoing wars for independence. The political pressure from Napoleon is therefore related to mass public education characteristics.

Using the political pressure from Napoleon as an instrument for the level of tracking in education systems leads to a strong first stage. When the number of tracks available to students is used as a tracking measure, it also leads to a significant second stage. Jointly interpreting the results on the number of tracks and the age of first selection, I find a positive effect of the level 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.

The main analyses use the mathematics test score to measure student performance. Similar results are found when using the science test score. For reading the results are less significant, signalling that tracking might only affect mathematics and science performance. Robustness checks are also done by varying the controls, the sample, and using nonlinear tracking measures. These checks show that including no controls leads to insignificant results, but controlling for either student or country controls leads to similar results, and that the sample selection has little influence on the conclusions. The nonlinear checks show that tracking seems predominantly positive for student performance when more than 2 tracks are available to students.

One important caveat of this analysis is that political pressure from Napoleon influenced many facets of European countries. The exclusion restriction is examined by looking at whether the political pressure from Napoleon is also related to current day societal characteristics. Although the relation between the political pressure from Napoleon and 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 is much weaker than the relation with tracking, the analyses suggest that at the very least pressure from Napoleon influenced (determinants of) the GINI coefficient and the expenditures on education which might have influenced educational opportunities and other aspects of education which have an effect on student performance. This could affect the results. However, including the GINI coefficient and the expenditures on education to the analyses hardly changes the results. Given these considerations, this paper provides an additional approach to supplement the existing evidence on the effects of education system characteristics on student performance.

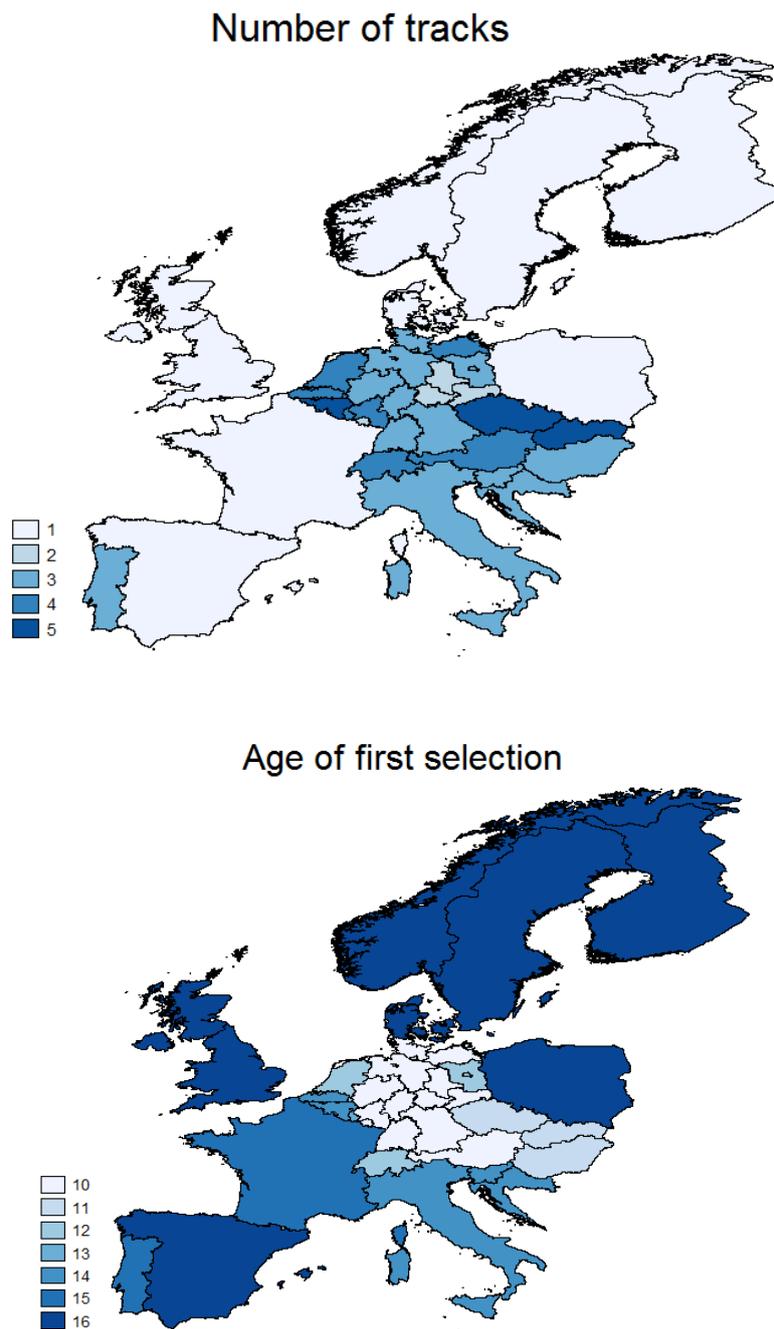
References

- Acemoglu, D., D. Cantoni, S. Johnson, and J. Robinson (2011) The Consequences of Radical Reform: The French Revolution. *American Economic Review*, 101, pp. 3286–3307.
- Archer, M. (1979) *Social Origins of Educational Systems*. SAGE Publications Ltd: London.
- Ariga, K. and G. Brunello (2007) Does Secondary School Tracking Affect Performance? Evidence from IALS. *IZA Discussion Paper*, no. 2643.
- Becker, S. and L. Woessmann (2009) Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *The Quarterly Journal of Economics*, 124(2), pp. 531-596.
- Boli, J., F. Ramirez and J. Meyer (1985) Explaining the Origins and Expansion of Mass Education. *Comparative Education Review*, 29(2), pp. 145-170.
- Cantoni, D. and N. Yuchtman (2014) Medieval Universities, Legal Institutions, and the Commercial Revolution. *The Quarterly Journal of Economics*, 129(2), pp. 823–887.
- Centennia Software (2002) *Centennia Historical Atlas: French Revolutionary and Napoleonic Era from 1789 to 1819*. www.historicalatlas.com
- Dittmar, J. (2011) Information Technology and Economic Change: The Impact of The Printing Press. *The Quarterly Journal of Economics*, 126(3), pp. 1133-1172.
- Elk, R. van, M. van der Steeg, and D. Webbink (2011) Does the Timing of Tracking Affect Higher Education Completion? *Economics of Education Review*, 30(5), pp.1009-1021
- Eurydice (2013) http://eacea.ec.europa.eu/education/eurydice/eurypedia_en.php
- GADM database of Global Administrative Areas (2013) *Codes and software for making maps in Stata*. www.gamd.org.
- Grab, A. (2003) *Napoleon and the Transformation of Europe*. Palgrave MacMillan: England.
- Hanushek, E. and L. Woessmann (2006) Does Educational Tracking Affect Performance and Inequality? Difference-in-Difference Evidence Across Countries. *Economic Journal*, 116, pp. C63-C76
- Horloge Parlante (2013) <http://www.horlogeparlante.com/nl/distance.php>.
- Kerr, S. T. Pekkarinen, and R. Uusitalo (2013) School Tracking and Development of Cognitive Skills. *Journal of Labor Economics*, 31(3), pp. 577-602.
- Korthals, R. (2016) The Education Revolution on Horseback I: The Relation between Napoleon Bonaparte and Education System Characteristics. *ROA Research Memorandum*, no. 2016/5.
- Kultusminister Konferenz (2013) *Übersicht über die Bildungsgänge und Schularten im Bereich der allgemeinen Bildung*. http://www.kmk.org/fileadmin/pdf/Bildung/AllgBildung/SchemaBildungsgaenge_und_Schularten-Stand_2012-12.pdf Last accessed 05-08-2013.

- OECD (2007) *Science competencies for tomorrows world*. OECD: Paris.
- OECD (2010) *School Factors Related to Quality and Equity. Results from PISA 2000*.
OECD: Paris.
- Ramirez, F. and J. Boli (1987) The Political Construction of Mass Schooling: European
Origins and Worldwide Institutionalization. *Sociology of Education*, 60, pp. 2-17.
- Woessmann, L. (2007) Fundamental Determinants of School Efficiency and Equity: German
States as a Microcosm for OECD Countries. *IZA Discussion Paper*, nr. 2880
- World Bank (2012) *World Development Indicators & Global Development Finance
Database*. Last accessed 23-03-2012. <http://databank.worldbank.org/data/home.aspx>
- Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. Pekrun, and R.
(2010): Programme for International Student Assessment 2006 (PISA 2006). IQB –
Institut zur Qualitätsentwicklung im Bildungswesen. Dataset.

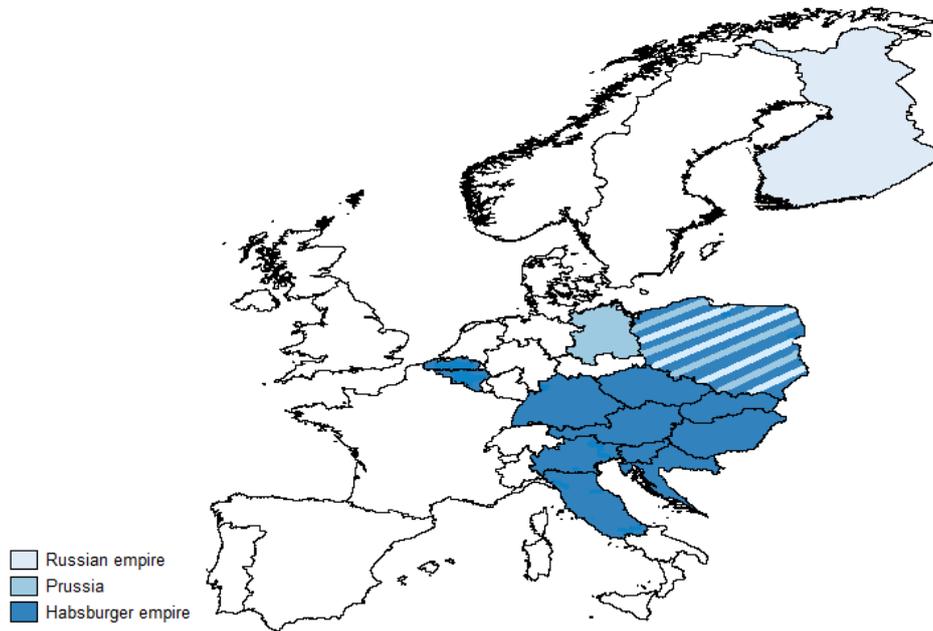
FIGURES

Figure 1: Education systems in Europe



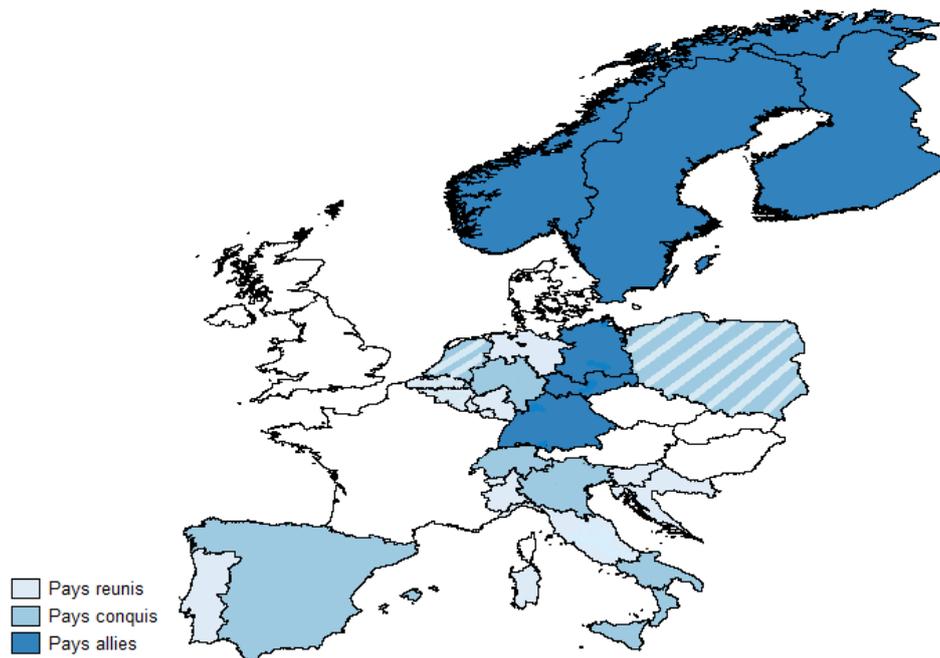
Notes: The borders are drawn at the lowest possible level at which the data are available. See Appendix A for more information.

Figure 2: Empires in Europe around 1800s



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

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



Notes: See main text for an explanation of the categories. Appendix A provides a justification of the borders used in this paper.

TABLES

Table 1: Descriptive statistics

	Number of tracks	Age of first selection	Distance	Empire	<i>Pays reunis</i>	<i>Pays conquis</i>	<i>Pays allies</i>	N
Austria	4	10	1033	1	0	0	0	4266
Bavaria (Ger.)	3	10	684	1	0	0	1	3634
Flanders (Bel.)	4	14	264	1	1	0	0	4759
Wallonia (Bel.)	5	14	264	1	1	0	0	2946
Croatia	3	14	1079	1	1	0	0	4575
Czech Republic	5	11	882	1	0	0	0	5783
Denmark	1	16	1026	0	0	0	0	4137
England-Wales	1	16	843	0	0	0	0	9877
Finland	1	16	2304	1	0	0	1	4621
Hanover (Ger.)	3	10	650	0	1	0	0	4283
Hungary	3	11	1244	1	0	0	0	4336
Kingdom of Naples (It.)	3	14	1291	0	0	1	0	5679
Luxembourg	4	13	287	0	1	0	0	2900
Netherlands	4	12	430	0	1	1	0	4317
Norway	1	16	1842	0	0	0	1	4291
North West of Italy (It.)	3	14	583	0	1	0	0	4355
Poland	1	16	1275	1	0	1	0	5437
Portugal	3	15	1452	0	1	0	0	4823
Prussia (Ger.)	3	10	878	1	0	0	1	7083
Rhineland (Ger.)	3	10	477	0	1	0	0	3066
Republic of Italy (It.)	3	14	639	1	0	1	0	9391
Saxony (Ger.)	2	10	849	0	0	0	1	3410
Scotland	1	16	843	0	0	0	0	2326
Slovak Republic	5	11	1088	1	0	0	0	4692
Slovenia	3	14	964	1	1	0	0	5778
Spain	1	16	1052	0	0	1	0	18579
Sweden	1	16	2044	0	0	0	1	3896
Switzerland	4	12	435	0	1	1	0	9501
Westphalia (Ger.)	3	10	578	0	0	1	0	4459

Notes: See text for sources.

Table 2: Correlations between age of first selection and test scores

Database:	PISA 2000	PISA 2006	PISA 2006	PISA 2006
Subject:	reading	reading	math	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 paper (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 paper (20 cnts)	0.19 ^b	0.19	-0.16	-0.11
Sample used in this paper (29 regions)	-	-0.14	-0.23	-0.35
<i>Notes:</i> ^a In PISA 2000 only NLD can be added. ^b This sample misses HRV, SVK and SVN.				

Table 3: The effect of tracking on mathematics test scores

Endogenous var: Model:	Number of tracks				Age of first selection			
	OLS	IV1	IV2	IV3	OLS	IV1	IV2	IV3
1st stage								
Distance		-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)		0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.001)
Empire			0.842*** (0.292)	0.960*** (0.283)			-0.609 (0.756)	-0.002 (0.731)
Pays réunis				0.667* (0.383)				0.562 (0.833)
Pays conquis				-0.015 (0.452)				0.438 (0.828)
Pays allies				-0.197 (0.347)				-2.055** (0.882)
2nd stage								
Number of tracks	9.04** (3.45)	11.86 (9.23)	16.45** (7.45)	17.00** (7.13)				
Age of first selection					-2.03 (1.83)	-7.92 (6.93)	-10.22 (6.90)	-6.43 (4.65)
# of students	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
R ²	0.33	0.33	0.33	0.32	0.32	0.31	0.29	0.31
Partial R ² 1 st stage	-	0.34	0.43	0.50	-	0.21	0.23	0.34
F test 1 st stage	-	63.69	28.09	16.25	-	24.70	10.72	10.04
<i>Notes:</i> The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. 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, 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, and GDP per capita 2005. Imputation dummies and interactions are included. Durbin and the Wu-Hausman tests are satisfied. Clustering on country level.								

Table 4: The relation between the political pressure from Napoleon and 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	0.08	0.03	0.11	-0.05
p F test	0.16	0.25	0.09	0.73

Notes: The superscripts *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See Appendix B for data sources.

Table 5: The effect of tracking on reading and science test scores

Model:	Number of tracks				Age of first selection			
	OLS	IV1	IV2	IV3	OLS	IV1	IV2	IV3
Mathematics	9.04**	11.86	16.45**	17.00**	-2.03	-7.92	-10.22	-6.43
Science	6.65*	10.10	14.70**	15.28**	-2.81	-6.75	-8.98	-6.6
Reading	4.81	6.69	11.27	13.17**	-0.61	-4.47	-6.54	-3.86

Notes: IV model 1 uses as instruments for the number of tracks the distance to Paris. IV model 2 uses distance and the empire dummy. IV model 3 includes also the Grab dummies. Each cell is a separate regression. See Table 3 for notes on the models.

Table 6: The effect of tracking on mathematics test scores: Results using different samples.

	Number of tracks				Age of first selection			
	OLS	IV1	IV2	IV3	OLS	IV1	IV2	IV3
29 countries (Main results)	9.04**	11.86 (63.69)	16.45** (28.09)	17.00** (16.25)	-2.03	-7.92 (24.7)	-10.22 (10.72)	-6.43 (10.04)
20 countries	9.31**	13.10 (55.05)	12.57* (16.65)	13.29** (9.93)	-5.21	-9.39 (56.61)	-7.99* (16.42)	-7.20 (6.83)
41 countries	6.59*	24.76*** (27.69)	24.85*** (18.07)	18.74*** (14.97)	-2.47	-15.94** (12.86)	-17.35** (6.50)	-13.66*** (3.63)

Notes: See Table 5.

Table 7: The effect of tracking on mathematics test scores: Results for the most extreme deviations when a single country is excluded

Country excluded	OLS	IV1	F	IV2	F	IV3	F	N
None (Main results)	9.04**	11.86	63.7	16.45**	28.1	17.00**	16.3	157,200
Flanders Belgium	7.17**	6.84	51.8	11.12**	24.6	11.71**	15.5	152,441
Kingdom of Naples, Italy	9.00**	9.13	61.3	13.15**	27.4	13.23**	20.8	151,521
Wallonia Belgium	9.01**	12.65	70.9	17.98**	24	18.77**	13.2	154,254
Luxembourg	10.90***	12.82	49.8	18.64**	25.1	20.09**	12.9	154,300
Finland	10.95***	21.32***	50.9	21.60***	19.6	20.56***	15.9	152,579
<i>Notes:</i> See Table 5.								

Table 8: The effect of tracking on mathematics scores: varying controls.

	Number of tracks				Age of first selection			
	OLS	IV1	IV2	IV3	OLS	IV1	IV2	IV3
Main results	9.04**	11.86	16.45**	17.00**	-2.03	-7.92	-10.22	-6.43
		(63.69)	(28.09)	(16.25)		(24.7)	(10.72)	(10.04)
Add edu. exp.	13.62***	20.96*	24.78***	22.47***	-4.70**	-17.08	-20.88**	-10.95**
		(19.13)	(11.84)	(9.50)		(16.66)	(6.95)	(7.17)
Add. gini	10.16**	12.92	16.31**	15.81**	-2.13	-7.07	-6.34	-4.55
		(36.25)	(19.62)	(14.07)		(32.42)	(36.89)	(19.53)
Add edu. exp. and gini	4.56	16.56*	14.43**	10.98*	-5.70**	-10.89*	-10.89*	-10.58**
		(19.52)	(11.83)	(12.16)		(26.72)	(15.66)	(13.78)
No controls	1.66	7.62	8.156	5.72	-2.83*	-5.40	-6.17	-8.71*
		(55.28)	(27.49)	(20.26)		(18.70)	(8.46)	(6.60)
Only student controls	11.43***	11.36	17.21**	17.31**	-2.92	-8.58	-12.96*	-8.31
		(55.00)	(28.24)	(18.18)		(15.60)	(7.63)	(5.55)
Only GDP control	1.770	5.36	8.75	6.25	-2.93*	-3.76	-5.62	-7.04*
		(70.53)	(33.29)	(20.30)		(24.89)	(10.10)	(12.62)
Only GDP and edu. exp.	5.17*	13.81	15.87*	9.87	-6.19***	-12.81	-15.83**	-12.91***
		(16.12)	(12.22)	(10.37)		(11.72)	(4.66)	(9.45)
Only GDP and gini	2.14	9.22	10.19	7.71	-1.51	-4.91	-4.74	-4.18
		(31.39)	(18.50)	(14.26)		(36.81)	(32.13)	(19.38)
Only GDP and edu. exp. and gini	4.56	16.56*	15.43**	10.98*	-5.70**	-10.89*	-10.89*	-10.58*
		(16.52)	(11.83)	(12.16)		(26.72)	(15.66)	(13.78)

Notes. See Table 3, except that these models contain no or different controls. Each cell is a separate regression. F test statistic of the first stage in parentheses.

Table 9: The nonlinear effect of the number of tracks and the age of first selection on mathematics scores

	OLS	IV1	IV2	IV3
More than 1 tracks	15.7	34.65 (69.37)	44.49 (28.23)	47.07* (11.97)
More than 2 tracks	11.88	36.19 (56.46)	49.61* (17.85)	48.21** (11.16)
Selection before 12	-2.06	85.6 (3.64)	118.1 (2.05)	3.87 (18.33)
Selection before 14	8.1	48.82 (8.00)	60.86 (3.80)	29.92 (5.22)
<i>Notes:</i> See Table 5.				

Appendix A- The borders used in this paper

The borders of the maps used in this paper are drawn by using the command `smap` 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 threat (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 paper Westphalia refers to the current German states Hesse and North Rhine-Westphalia. The Rhineland is 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.¹² Schleswig-Holstein belonged to Denmark around 1800 and is therefore display as a single territory in the historical.

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 paper 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,

¹² 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 paper.

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.

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 B- The exclusion restriction

Table 4 in the main text shows the relation between the political pressure from Napoleon and a number of societal characteristics, namely 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 it is not unreasonable to think that Napoleon has had an influence. The GINI index is the most recent number (2008-1999) from the World Bank. The number of police officers per 100 000 inhabitants for 2008 is taken from Eurostat. The expenditures on education are from the World Bank (Public spending on education, total (% of GDP)) for 2005. The proportion of seats held by women in national parliaments for 2008 is from the World Bank. Table 4 in the main text displays results from the main IV model (IV model 2). Tables B1 and B2 in this Appendix display similar analyses using IV model 1 (using distance) and IV model 3 (using distance, empire, and the three Grab dummies).

In Table 4 it can be seen that the expenditures on education seem to have a significant relation with the political pressure from Napoleon and the F test statistics is below 10% significance level ($p=0.087$). In Tables B1 and B2 it can be seen that the expenditures on education also has a significant relation with the political pressure from Napoleon in IV model 2. The GINI coefficient has a significant relation in both IV model 1 and IV model 3. 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 and the expenditures on education, this exclusion restriction is violated.

Table B1: The relation between the political pressure from Napoleon and societal characteristics: IV model 1

	GINI	Police	Exp. Edu.	Women seats
Distance	-0.003* (0.001)	-0.050 (0.033)	0.000** (0.000)	0.003 (0.004)
Constant	33.550*** (1.343)	364.500*** (35.370)	4.359*** (0.398)	24.570*** (3.877)
# of countries	26	29	29	29
Adjusted R ²	0.11	0.04	0.12	-0.01

p F test	0.05	0.14	0.03	0.44
<i>Notes:</i> 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.				

Table B2: The relation between the political pressure from Napoleon and societal characteristics: IV model 3

	GINI	Police	Exp. Edu.	Women seats
Distance	-0.001 (0.001)	-0.006 (0.041)	0.001* -0.001	-0.001 (0.004)
Empire	0.123 (1.196)	30.54 (33.56)	-0.295 (0.411)	-0.249 (3.53)
Pays réunis	-0.937 (1.56)	14.92 (41.26)	0.12 (0.505)	4.448 (4.34)
Pays conquis	1.402 (1.496)	-0.979 (41.96)	-0.257 (0.514)	2.691 (4.414)
Pays allies	-4.630** (1.793)	-94.330* (49.62)	-0.004 (0.608)	14.260** (5.219)
Constant	32.980*** (2.069)	323.000*** (56.24)	4.467*** (0.689)	22.590*** (5.915)
# of countries	26	29	29	29
Adjusted R ²	0.32	0.09	0.01	0.11
p F test	0.02	0.22	0.41	0.18
<i>Notes:</i> 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.				