Central School Exams and Adult Skills: Evidence from PIAAC

Lisa Leschnig* Guido Schwerdt* Katarina Zigova*

Preliminary and incomplete draft

May 19, 2016

(Please don't quote or circulate.)

Abstract

Centralized exit exams are often hypothesized to favorably affect incentive structures in schools. Previous research indeed provides vast evidence on positive effects of centralized exams on student test scores, but critics warn that these effects may arise through strategic behavior of students and teachers, which may not affect human capital accumulation in the long-run. Exploiting variation in examination types across school systems and over time, we document for the first time that central exams are also associated with higher adult skills. This finding is in line with true productivity-enhancing effects of central exit exams.

JEL-Code: I20, J24, J31

Keywords: Central exit exams, Adults skills, PIAAC

We thank the participants of the Public Economics Seminar at the University of Konstanz and of the Spring Meeting of Young Economists in Lisbon for helpful comments. Any errors are our own.

*University of Konstanz, Department of Economics

1 Introduction

Central exit examinations (CEE) are associated with substantially higher test outcomes of students (see Bishop 1997; Woessmann 2003, 2005; Jürges et al. 2005; Fuchs and Woessmann 2007). While this reduced-form pattern is well documented, critics warn that this result may simply reflect differences in students' test-taking ability, rather than actual differences in knowledge and skills (Popham 2001; Volante 2004). If that were so, central exams would not genuinely improve human capital that affects productive skills of adults. Indeed, the existing evidence on the relationship between central exams and labor market outcomes is at best mixed (Backes-Gellner and Veen 2008; Piopiunik et al. 2013) and, to our knowledge, no study has yet investigated the link between central exams and cognitive skills of adults.

Our paper closes this gap in the literature by investigating the relationship between the type of exit examination at the end of secondary school and cognitive skills during adulthood as measured in the recent OECD study "Programme for the International Assessment of Adult Competencies" (PIAAC 2013). The PIAAC data allows a comprehensive analysis of the relationship between CEEs and cognitive skills as well as labor market outcomes. We supplement the PIAAC data with specifically collected data on the type of exit examinations for 23 participating countries in PIAAC over the last 50 years. This allows us to study the potential impact of central exams on adult skills by exploiting the cross-country and within-country variation in exam types over time. If the widely documented positive relationship between CEEs and learning outcomes of students is a mere side effect of test-taking ability, we would not expect to see any systematic association between the type of the examination in secondary schools and cognitive skills of the adult population.

Our findings suggest a substantial positive effect of central exit exams on adult cognitive skills. Conditional on covariates, individuals in education systems with centralized exams at the end of high school are associated with more than 20 percent of a standard deviation higher skills during adulthood. Taking unobserved country variation into account, CEE graduates still significantly outperform graduates of schools using local exams by more than 6 percent of a standard deviation in numeracy and by more than 12 percent in literacy and problem solving scores. Our results are robust to a large set of alternative specifications. In particular, we show that a CEE regime significantly increases the probability of employment and higher educational attainment. Exploiting variation in exam types across German federal states, we document CEE effects of similar magnitudes even within one country. Lastly, we estimate the CEE effects using data from the 2000 wave of the “Programme for International Student Assessment” (PISA) along the lines of the analysis by Fuchs and Woessmann (2007). Limiting the PIAAC data to the 2000 PISA cohort, we estimated in both datasets CEE effects of roughly the same magnitudes. In sum, our findings support the hypothesis that centralized exit exams have long term productivity-enhancing effects.

Our empirical study leans on previous theoretical results on the format of screening to students' learning. Becker's (1982) seminal publication highlights the role of screening accuracy. From Becker's model it follows that the increase in the precision of the screening lead to the increase in student's learning, at least for students with the above modal knowledge. In the subsequent model Becker and Rosen (1992) show that the learning effect is the highest if high-school students have to compete to achieve higher grades, to distinguish them from their peers, rather than only obtain a pass-fail high school diploma. Moreover Becker and Rosen (1992) model documents the enhancing effect of an external standard, known ex-ante, to students' effort. These two models offer theoretical grounds for why the external standards are relevant for student learning. Bishop and Wößmann (2004) offer a signaling explanation of students effort. Students under central exams learn more, because the diploma obtained has reliable signaling value of their ability and productivity and moreover, because the peer pressure effect of low performers decreases.

Empirical literature already evidenced some of these theoretical predictions using national or international survey data. Already in 1997, Bishop documented significant differences between students test outcomes in mathematics and science of 13-year-olds across 39 countries with or without CEE regimes. Employing the “Third International Mathematics and Science Study” (TIMSS) data, Bishop (1997) quantified the CEE impact in mathematics between one and two U.S. grade level equivalents. CEE effects of a similar magnitude were found few years later by Woessmann (2002) using international TIMSS and TIMSS-repeat data. Woessmann estimated the CEE effect to be about one year of schooling. In his two consequent papers Woessmann (2003; 2005) studies the international differences in TIMSS and PISA data focusing on all possible institutional variables affecting student learning. In Woessmann (2003) he finds that even taking a broad spectrum of institutional variables into estimation, the effect of CEEs is significant of about 0.16 of the international standard deviation. In Woessmann (2005) he estimates interaction parameters with CEEs instead. The interaction terms with CEEs were strongly positive for school autonomy and school budgeting variables, highlighting the role of CEEs as a tool for school accountability. The CEE interaction terms with parental low education and parent born abroad dummy were positive and significant, too. Hence CEEs affect stronger disadvantaged students, which contributes to higher equality in education. All these studies are using only cross country variation, thus the estimated CEE impact might be mixed with other international differences in education systems. To estimate a causal effect of CEEs, Jürges et al. (2005) exploit regional variation in CEE regimes in science and math across German federal states. Using a diff-in-diff estimator applied to TIMSS data for Germany, they find a significant increase in students achievement due to the existence of CEEs, of about one-third of a school year, or 0.16 of a national standard deviation. Finally, Fuchs and Woessmann (2007) show significant improvements in math, science and reading test results of 15-year-olds students in countries where CEE system is in place, using the international PISA 2000 data.

From the labor market perspective it is important to assess whether the documented improvement in students' test results under the CEE regime is followed by higher employment probability or higher wage. Here is the evidence mixed and limited to Germany. Backes-Gellner and Veen (2008) find no wage premium for high-school graduates from the CEE federal states and explain this with the existence of the pooled labor market in Germany. On the contrary, the more recent study by Piopiunik et al. (2013) finds that central exam graduates enjoy higher earnings and lower unemployment if they enter labor market shortly after obtaining a high school diploma. Our study contributes to this existing evidence by thoroughly documenting a long-term impact of central exams on adult skills and labor market outcomes.

The next Section summarizes the largely self collected dataset that details the use of CEEs in the OECD countries which participated in PIAAC. Section 3 describes the PIAAC data structure and presents descriptive statistics. In Section 4 we explain our empirical approach. Section 5 presents our estimates of the CEE effect in baseline and auxiliary specifications. The last Section concludes.

2 Central Exit Examinations across OECD Countries

A central school leaving examination is a written test, administered by a central authority (e.g. ministry of education) which provides centrally developed and curriculum based test questions, covering core subjects such as math and the first language. CEEs set exogenous quality standards and therefore ensure comparability of students' performances/skills through a consistent test implementation and grading as well as a passing threshold (Backes-Gellner and Veen 2008). CEEs must be delimited from commercially prepared achievement tests such as SAT[1] or California Achievement Test (CAT) scores. CEEs are required to have direct consequences for students passing them, e.g. by representing a significant part of the final grade. Commercially prepared tests or university entrance exams do not feature these direct consequences or are not taken by a sufficient large number of graduates and consequentially are not considered as a CEE (Bishop 1998, Bishop 1997). CEEs can be organized either on a national level or on a regional level and must be mandatory for all or at least the majority of a cohort in a certain school track.[2]

Table A1 reports the CEE introduction years across the 23 countries of the PIACC dataset used in our analysis. Out of them Belgium, Austria and Spain are currently the only non-CEE countries. In Austria the introduction of a central school leaving exam took place in 2015. Sweden is an exception among the PIAAC countries as their CEEs were eliminated in 1968. Since then, according to our knowledge, Swedish students can pass an centralized examination but on a voluntary basis. Given that this CEE is not mandatory, it does not meet our definition and we treat the Swedish system as non-CEE for the years after 1968.[3] The tradition of CEE started with France, where a central testing of high school graduates was already introduced at the beginning of 19th century. Denmark followed hundred years later in 1908. In 1919 and 1923 also Finnland and Italy switched from a local to central examination system. Ireland, UK, Norway, the Netherlands, but also Japan and Korea followed sequentially during the second half of 20th century. The last large wave of countries switching to CEEs took place in Eastern Europe at the turn of the 21th century with Estland in 1997, Poland and Slovak Republic in 2005, Cyprus in 2006 and finally the Czech Republic in 2011. Figure 1 maps the dynamic variation in CEE regimes across PIAAC countries (except for Germany) between 1960 and 2012.

In Australia, Canada, Germany and the United States, the school system is organized on a federal state or a province level. This means that educational policy is not consistent and the decision if and when a CEE is introduced varies within the respective countries. Table A2 gives more detailed information about the examination systems which are in place across the federal states and provinces of the four decentralized countries.

[1] Abbreviation SAT has no unambiguous meaning. Former known as Scholastic Assessment Test, Scholastic Aptitude Test or Scholastic Achievement Test

[2] See Bishop (1997, pp. 260) for a list of further requirements not explicitly applied in our definition as they are automatically fulfilled by all CEE-countries meeting our definition.

[3] Deviating from us, Fuchs and Woessmann (2007) assume for Sweden a share of 50% of students participating in CEEs in the year 2000. Bishop (1998) assumes no CEEs for Sweden.

In Australia, currently all of the states have introduced the CEEs. The state with the longest tradition in central testing of graduates is New South Wales which introduced CEEs in 1967. For the states South Australia, the Northern Territories, Tasmania, Victoria and Western Australia we know that they had CEEs in 1992 and so they must have introduced central testing sometime before. Finally, Queensland switched to CEEs in 2008. There is also substantial heterogeneity in the presence of CEEs within the Canadian provinces, too. Currently, six of its thirteen provinces have centralized school leaving examinations of which New Brunswick is further divided with respect to CEEs. The francophone part of the New Brunswick introduced CEEs in 1991, while the anglophone part did not—until today. Three provinces eliminated CEEs in 1960' and 1970's, out of which Manitoba reintroduced CEEs in 1991. In the USA, no state but one applies central examinations. The state of New York introduced CEE in 2003. The remaining states apply different standardized tests that do not meet our definition of central exit examinations. For a detailed survey concerning high school graduation testing across US states see Caves and Balestra (2014).[4]

In Germany the CEE system currently clearly prevails. As of 2012 all but one state of Rhineland-Palatine had CEEs in place on the upper secondary level. However prior the 1990's most states had only the local school leaving examinations in place. The three exceptions are the southern states Baden-Würtemberg, Bavaria, and Saarland which hold a long tradition in the central testing of high-school students. This has changed during 1990s when Mecklenburg-Western Pommerania, Thuringia and later Saxony and Saxony-Anhalt switched to CEEs. The last big wave of switching states took place in the 2000s with Brandenburg, Hamburg, Berlin, Lower Saxony, North-Rhine Westphalia, Bremen, Hesse and finally Schleswig-Holstein. Figure 1 summarizes the dynamic variation in CEE regimes across the German states between 1960 and 2012.[5]

Unfortunately the PIAAC data only contain information on a federal or provincial base for Germany. This is a problem for the remaining counties having a decentralized school system as we cannot determine whether an individual, graduated in Australia, Canada or the USA in a certain year, was exposed to a CEE regime or not. We solved this by conducting a dynamic CEE factor between zero and one according to Fuchs and Woessmann (2007) but with our self-collected regional data for Australia,

[4] According to Fuchs and Woessmann (2007), the share of US high-school students under the CEEs regime was 0.1 in 2000.

[5] For a detailed summary of the German school system see e.g. Jürges et al. (2005), Schwerdt and Woessmann (2015).

the USA and for anglophone and francophone Canada for the years between 1960 and 2012. For this purpose, we weight a regional CEE dummy which can be either 0 or 1 for each federal state and each year with the state's respective population share and then aggregate over the provinces/federal states. The resulting factor amounts to the probability that an graduate from a specific year and country passed a centralized exit examination. The dynamic CEE factor for Australia ranges between 0 and 1 and for the USA between 0 and 0.063. For Canada we conduct two factors, each for the anglophone and francophone part. The CEE factor of the English speaking part of Canada ranges between 0 in 1972/1973 and 0.623 in the 1960s while we calculate a less volatile factor for the French Part of about 0.946 to 1. Figure 2 maps the variation of CEE factors for Australia, the USA and (English and French) Canada over time.

3 PIAAC Data

The second source of data, we employ in our analysis, are survey data on adult skills in numeracy, literacy and problem solving skills in technology rich environments. These data stem from the unique survey performed in 24 countries within the OECD "Programme of International Assessment of Adult Competencies" (PIAAC, 2013). The PIAAC survey builds on, and significantly extends, the earlier OECD skills surveys: "International Adult Literacy Survey" (IALS) and "Adult Literacy and Lifeskills Survey" (ALL).[6]

The PIAAC survey focuses on measuring numeracy, literacy and problem solving skills in nationally representative sample of active age population between 16 and 65. The survey took place between August 2011 and December 2012. The participating countries were Australia, Austria, Belgium (only Flanders), Canada, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, the Netherlands, Norway, Poland, the Russian Federation, the Slovak Republic, Spain, Sweden, the United Kingdom (only England and Northern Ireland), and the USA.[7] The countries had varying sample sizes, from 4.000 to 10.000, and different sampling strategies.[8] This was of no harm to the pooled data, as each individual was endowed

[6]IALS represented a first attempt to assess the literacy skills of entire adult populations in a framework that provided data comparable across cultures and languages. IALS was performed in 20 countries in three rounds over 1994-98. ALL survey measured the literacy and numeracy skills of a nationally representative sample of 16- to 65-year olds in ten countries in two rounds over 2003-08.

[7]Nine more countries: Chile, Greece, Indonesia, Israel, Lithuania, New Zealand, Singapore, Slovenia, and Turkey join the 2nd round of the PIAAC survey in 2016.

[8]Most countries had a sample size of around 5.000, what was the OECD target sample size. Canada had an exceptionally large sample of 26.000 individuals.

with a post-sampling weight representing the exact share of his socio-economic strata in the population of the country. In our baseline estimations we use data from all countries except Russian Federation.[9] In our empirical analysis, we standardized the PIAAC measures of skills from the original 500-point scale to have a global mean within each skill category equals to zero and a standard deviation of one. The standardization allows for interpretation of the estimated coefficients in terms of standard deviations. In addition to assessing adult skills, the PIAAC collects background data on individual labor-market status, earnings, education, experience and numerous demographic characteristics.

To assess whether an individual finished high-school in a country with a CEE regime, we joined the self-collected information on the CEE introduction (Section 2) and an individual's year of secondary school graduation. Note that an individual was assigned to a CEE regime, even if she might never participated in such an exam, just by the fact that the country where she graduated, from any type of secondary school, had at the time of her graduation a CEE practice. We do this in order to measure a more general effect of CEEs on the quality of schooling and consequently on its persistence in adult skills. The CEE variable is for most of the PIAAC individuals a 0/1 variable. Australia, Canada and the USA are countries with a federal school system, hence the introduction of the CEE lies in the hands of its states. Ideally we would like to assign a 0 or 1 to each individual according to the state of her graduation. But the PIAAC data in these three countries do not contain a state indicator. Due to this data deficiency in these three countries, we assign to each individual a so-called CEE factor, i.e. a number between 0 and 1, instead of a CEE dummy.[10] The CEE factor is calculated as a share of states, weighted by population, which has the CEE in place and thus reflects a probability that an individual in such country was exposed to a school system with CEEs. The development of the CEE factor in these three countries over time is depicted in Figure 2.

Table 1 surveys the means or shares of variables we use in our regressions, distinguished by the CEE setting.[11] The means of standardized skills scores are significantly higher for individuals graduated in school systems with CEE than those without it. However this can only be an age effect. The age of individuals in CEE-states is on

[9] Data from Russian Federation are not representative of the entire Russian population because they do not include the population of the Moscow municipal region.

[10] Germany is a country with a federal school system, too. But the German PIAAC dataset contains the state indicator.

[11] Individuals from Australia, Canada and the USA are included in the CEE (non-CEE) column if their CEE factor is > 0.5 (< 0.5).

average about 6 years lower than in non-CEE states. This is a natural artifact of the gradual introduction of CEEs during the second half of the 20th century (see Section 2). Hence the younger individuals have a higher probability to live in a CEE state. Age composition effects are probably causing the small positive gain in earnings of the non-CEE individuals due to their seniority and on the contrary higher employment share of the CEE individuals due to the fact that comparatively younger individuals (< 50) are more probable employed. The ISCEDs [12] variables are dummies which bear 1 if an individual's attainment is at least ISCED 3 (5). While the share of the tertiary attainment is the same, the CEE individuals have a little lower ISCED 3+ attainment, which we explain by differences in the school systems of the CEE vs. non-CEE countries, as we do for the differences in the educational attainment of the parents. The gender composition of the CEE and non-CEE individuals is the same. There are small differences in the age at graduation and the second-generation migrant status.

4 Empirical Strategy

However, up to this point, the current evaluation of CEEs is not unanimous in the academic literature. On the one hand, central school exams have been argued to increase accountability in schools systems, which may have real positive effects on student learning. It is already well documented that CEEs are associated with substantially higher learning outcomes of students (See Bishop 1997, 2006; Jürges et al. 2005, Woessmann 2003, 2005). On the other hand the existing evidence on the relationship between central exams and labor market outcomes is at best mixed (Backes-Gellner and Veen 2008; Piopiunik et al. 2013). To our knowledge, no study has yet investigated the link between central exams and cognitive skills of adults. Moreover, many researchers argue that the intended purpose of the CEEs on students performance is often blurred due to the effort of their instructors to drill only the students' test competencies and expected test material, rather than teaching them the whole scope of the curricula. This is the so-called "teaching to the test" effect studied mainly by education scientists. They maintain that high-stake exams for both students and teachers will lead to a decrease in the quality of teaching (Popham 2001, Volante 2004).

The intention of this study is to properly asses the long term effects of the CEE regime to adult competencies. Because if the widely documented positive relationship between CEEs and learning outcomes of students is a mere side effect of "teaching to

[12] ISCED stands for "International Standard Classification of Education". In our analysis we use the OECD categorization of ISCED categories from 1997.

the test", we would not expect to see any systematic association between the type of the examination system in secondary school and cognitive skills of the adult population.

From this perspective it is not important whether an individual literally passed the CEE but rather whether she was exposed to a school system with a centralized assessment at the time point of her high school graduation. We achieve the long term effects due to availability of independent cross-sections with information on competencies of the a prime-age population, aged 16 to 65 years, across 23 OECD countries between 1960 and 2012. Combined with changes in institutional settings between 1960 and 2012 within the OECD 23 countries, we have enough within and across country variation in the CEE regime induced by the data's pseudo panel structure. Thus, even if our data miss the randomized experiment features, and we cannot claim any clean causality from CEE regimes to competencies, we are confident to estimate reliable associations.

We estimate an OLS model with the individual standardized competency score $s_{i,c,g}$ as a dependent variable

$$s_{i,c,g} = \Delta_c + \theta_g + \gamma\, CEE_{i,g} + X_i\beta + \varepsilon_i; \qquad \text{where } i = 1, \ldots, N. \tag{1}$$

The right hand side variables includes the country fixed effects Δ_c as well as age fixed effects θ_g. The variable of interest, $CEE_{i,g}$, equals one if an individual, graduated in country or region c in graduation year g, was exposed to CEE settings during his high school graduation and equals zero otherwise. For Australia, Canada and the USA this variable takes values between 0 and 1, representing a fraction of CEE graduates. We expect a positive CEE effect to competency, which is fading out with age, but remains always a positive difference in skills between individuals in CEE and non CEE regimes. This CEE effect is thus measured by the parameter γ which, due to the standardized numeracy skill measure $s_{i,c,g}$, expresses changes in standard deviations of s. The matrix X_i contains solely exogenous personal and family background variables and a dummy on the individual's age at graduation which is used as an additional fixed effect in some specifications. In addition to age and age at graduation fixed effects, we include dummies on gender, the migration background and dummies for parent's attained education. Furthermore, we assume an error term ε_i which is clustered on the level of 39 CEE regions crossed by 47 different graduation cohorts, which is the finest grid on which the examination systems in our data can vary. Finally, N is the number of sample individuals. As we employ country fixed effects, we identify the CEE effect by using only within country variation in exam regimes over time.

In the baseline estimations of the model 1, we focus on estimation of the CEE effect in numeracy skills, as a competency in numeracy is easier to compare across different countries with different languages. In the subsequent estimation rounds we also test for the presence of the CEE effects in the other two competencies, namely numeracy and problem solving as well as labor market outcomes and education attainment.

5 CEE Estimation Results

We present our estimation results of CEE effect in the following sequence. The baseline estimates consist of 6 variations of specification (equation 1) to assess the magnitude range of the CEE effect in numeracy skills. To confirm the robustness of the estimated effect of CEEs in numeracy, we estimate the CEE effect in three sets of auxiliary specifications. In the first set, we test the presence of the CEE effect in other outcomes, namely in the remaining two competencies, labor market outcomes and probability of the higher education attainment. The second set limits the PIAAC sample to Germany. Due to the existence of the within country variation in the CEEs regimes of German states, we test here the internal validity of the previous results. Finally in the last set, we estimate the CEE effects in two datasets of students' test outcomes in math, namely PISA 2000 and TIMSS 1999 and compare them with an estimated CEE effect within an equally abridged PIAAC sample to attempt for an external validation of the previous CEE results in numeracy.

5.1 CEE Effects in Numeracy Skills

We survey our baseline estimates of the CEE effects in numeracy skills within PIAAC countries in Table 2. The columns (1) to (6) of Table 2 present the CEE effect across models with increasingly complex specification. Conditional only on covariates, individuals in education systems with CEEs at the end of high school are associated with more than 20 percent of a standard deviation higher skills during adulthood (see columns (1) to (3) of Table 2). Taking unobserved country variation and flexible age specification via fixed effects into account, column (4) to (6), graduates in CEE regimes still significantly outperform graduates in local exam systems by 6 to 8 percent. Given that the dependent variable is standardized numeracy score, the point estimate of the CEE variable measures the percent increases in standard deviation of the numeracy. We treat specification (6) as our benchmark specification, due to the flexible treatment of age. We repeatedly estimate the specification (6) in other settings below.

Even if parents' education remains the overwhelming predictor of the individual numeracy scores, all specifications demonstrate a positive and significant effect of CEE regime in education systems over the PIAAC countries. All baseline specifications (Table 2) employ the sampling weights. We show later the stability of the CEE effects under alternative set of weights, and absence of weights, respectively. In each specification of Table 2, we calculate robust clustered standard errors, where the clusters are made out of crossing the 39 CEE regions with 48 graduation cohorts.

It is not straightforward to compare our results to other international studies on the CEE effects, as these mainly focus on the students' test differences in math and reading. Our data deviates in two important aspects: (i) they refer to numeracy and literacy skills, rather than school competency in math and reading; (ii) PIAAC encompasses the whole adult age population, rather than students of single age or grade cohort. Nevertheless, the CEE effects in skills demonstrated by our study and the CEE effects in tests results shown by previous empirical studies should not be qualitatively different. Fuchs and Woessmann (2007) find that participated PISA students under the CEE regime attain 19.5% of the standard deviation higher math test results than students from school systems without it.[13] Woessmann in his earlier study (Woessmann, 2002) estimates CEE effects using TIMSS and TIMSS-repeat data find a significant CEE effects in math ranging from 35 to 47% of an international standard deviation in test scores. Finally Bishop (1997), using the TIMSS data estimates 23 percentage points increase in correct answers for an average student of a CEE country.

5.2 CEE Effects in Other Outcomes

We are also interested in whether the CEE regime leads to better results in other skills, or moreover to better labor market outcomes and higher education attainment. We survey the estimated CEE effects in other outcomes in Table 3. We use the benchmark specification (Table 2, column 6), but replace the numeracy with the respective outcome.

The CEE effect in literacy and problem solving competencies (columns 1 and 2) do not qualitatively differ, however they are about twice as large than the estimated effect in numeracy, reaching 12 and 14% of the international standard deviation in literacy and problem solving. The parents' education has similar strong effect, but the gender gap disappears in literacy skills. We explain the stronger CEE effect in problem solving skills as a compensation for larger age fixed effects. Similarly, the larger CEE effect in literacy we explain by larger cross country variation in literacy scores, hence again the

[13] Fuchs and Woessmann (2007) work with PISA 2000 points directly and report 19.5 more points, but given that PISA data have international mean of 500 points and international standard deviation of 100 points, we can use directly the points effect as percentage.

large CEE effect partly compensates for this.

In addition, we use the PIAAC data to test the hypothesis that the existence of CEEs at the end of high school is associated with higher returns to skills for labor market entrants. Such relationship can be expected if CEEs indeed increase the signaling value of educational credentials and grades (see Schwerdt and Woessmann 2015). The probability of employment indeed slightly increases for individuals having the CEE school certificate (column 4). But the positive effect is insignificant in earnings (column 3).

Moreover our results in Table 3 give support to the view that CEEs increase probability of tertiary education attainment (see column 6), reflecting the fact that individuals from CEE states are better prepared for higher education and are more prone to finish it. The positive effects of CEE in employment and in tertiary educational attainment signal another argument for CEEs as an appropriate school quality tool with positive consequences going beyond the sole school test results.

5.3 CEE Effects in German Subsample

Four of our 23 PIAAC countries are characterized by a decentralized school system where the CEE regime is set by state, rather than by the federal government. We make use of this within country variation in Germany by using the PIAAC information on the individual's state of residence.[14] We repeat our baseline estimation with the subsample of German individuals, for all three competency measures, the two labor market outcomes and the two attainment dummies to make our results comparable to other studies based on German data and to test the internal validity of the previous results. Table 4 summarizes the estimates for all outcomes based on the German subsample.

All background variables mostly have the expected effects on the three competency measures and labor market outcomes. The level of parents' education again explains the largest part of the variation in skills and also significantly affects individual earnings. Graduates from the former GDR states perform worse in most outcomes, compared to those from the western federal states. However we excluded individuals from the former GDR with a graduation year before 1991, so that the indicator captures persistent historical differences in East German schooling and labor markets that are independent of CEE regime. Being female has a significantly negative effect in all outcomes except for literacy.

[14] Australia, Canada, and USA are the remaining three countries with decentralized school systems, but the PIAAC scientific use files does not include information on the individual's state of residence.

The CEE effects for the German subsample remain positive and significant throughout most of the specifications, but are slightly smaller than the general effects in Table 2 and 3. For numeracy the German CEE effect is 11% of the standard deviation. This effect is in line with the results of Jürges et al. (2005) with German TIMSS micro data. They apply a difference-in-difference by subject estimate by regressing the difference in skills between the CEE subject math and the non-CEE subject science on a set of background variables.

The estimated CEE effect on earnings is not statistically significant, which corroborates a previous German study by Backes-Gellner and Veen (2008) which also do not find any effect of CEEs. Contrary to our result, Piopiunik et al. (2013) find a positive and significant effect of CEEs on earnings, albeit only for students leaving schools directly for labor market. As we cannot control for the time span between graduation and labor market entrance, our results could be downward biased. On the other hand, according to Mincerian equation, the most important factors for labor earnings are the level of skills and experience, as shown recently also with PIAAC data by Hanushek et al. (2015). In our specifications we do not control for variables which can be potentially endogenous to skills, hence our specification explaining the earning could well suffer from omitted variable bias. The probability to be employed increases by 2.8 % for individuals from CEEs school systems and is statistically significant. This result supports the previous finding by Piopiunik et al. (2013) who estimate a negative effect of CEEs on unemployment of about the same size using SOEP data.

5.4 CEE Effects in PISA and TIMSS vs. PIAAC

In the previous sections we find a positive effect of CEEs on adult skills. It is now interestingto see, whether the baseline specification, used in the previous analysis, gives quantitatively similar results in well-known data sets on student learning outcomes. Both Woessmann (2002) and Jürges et al. (2005) investigate the role of CEEs in the education production with TIMSS Repeat and TIMSS 95 data, respectively. Woessmann (2002) finds a robust and positive effect in math scores of about 40% of a standard deviation using a cross-section analysis, while Jürges et al. (2005) find an effect of about 13 % of a standard deviation applying a difference-in-differences approach. Fuchs and Woessmann (2007) also apply a cross-section approach with PISA 2000 data and also find a lead of 19 PISA points for individuals exposed to a CEE regime. A simple direct comparison of the single results is not possible as our estimate contains country and age effects.

In order to test the plausibility of our PIAAC results we limited the PIAAC data to match the PISA 2000 and the TIMSS Repeat sample. In the first case, we limited the PIAAC sample only to those who have been 15 years old in 2000. IN the second case, we restricted the PIACC sample to individuals bron in the years 1984 and 1985. In both cases, we further limited PIAAC data to common participating countries of PISA and TIMSS.

Table 5 shows the results of our baseline specification in the restricted PIAAC data and PISA 2000 and TIMSS Repeat data sets. This is the specification (3) in Table 2 as we cannot include fixed age effects due to the cross-sectional structure of the data and country fixed effects due to identification problems. Colums (4), and (5) of table 5 are based on sample of PIAAC individuals who were 15 years old in 2000 (i.e. match to PISA), while columns (2) and (3) are based on sample of PIAAC individuals who were born in 1984-85 (i.e. match to TIMSS). The columns (3), (5), and (7) impose the CEE coding adopted from Fuchs and Woessmann (2007) (see A3), while the remaining columns take our CEE coding introduced in Section 2. Column (1) repeats column (4) but including all PIAAC countries.

The estimates for the 15 years old are larger, than the general PIAAC results (35 to 42%, vs. 27%), which we can be explain by the fact that the PISA 2000 equivalent cohort was about 25 years old in 2011/12 when the PIAAC data were collected. Younger individuals achieved comparable higher skill scores. All background variables show the expected signs and significantly explain the variation in math scores and numeracy. The Eastern European dummy grasps the differences in the different coding of CEEs. Again, the parents' education accounts for the major part in the outcome variation. Fuchs and Woessmann (2007) report a positive effect of central exams on math scores of 19.5% of a standard deviation. This is about half of the effect we estimate with our PISA estimates (35.6 %). Contrary to our estimations, Fuchs and Woessmann (2007) include a broad range of control variables and 32 countries. Together with the age effect this explains the comparatively large CEE estimates. The results in the TIMMS Repeat are in a similar range from 37 to 41 % of a standard deviation, depending on the CEE coding. This is in line with the estimates of Woessmann (2002) who finds 40% of a standard deviation but exceeds the 13% estimate of Jürges et al. (2005). The latter is no surprise as Jürges et al. (2005) estimate a clean causal effect with German results. The PIAAC estimates of the CEE effect of the TIMSS-like restriction (column 2) exceeds the result of the PISA-like restriction (column 4) with 51% to 41% of a standard deviation, applying our CEE coding. This also the case for the CEE coding of Fuchs and Woessmann (2007) but the difference between the to effects in the different restrictions is smaller (53%, column 3 to 48%, column 5). It is worth to note, that PIAAC tests nummeracy skills while PISA and TIMMS uses tests on math skills.

In all cases, the PIAAC CEE estimates provides higher values compared to PISA and TIMSS data. All results using the Fuchs and Woessmann (2007) coding yield higher estimates because Eastern Europe is indicated as 1, hence it compensates the negative estimate.

5.5 Robustness Checks

We scrutinized the validity of our baseline CEE results (Table 2) in three sets of robustness checks. In the first set we test for stability of the CEE effect of the benchmark specification by piecewise dropping countries from our sample. Table 6 documents the CEE effects resulting of this exercise. The CEE impact remains significant in all cases and range between and between 4.1% to 12.6%, where most of the country deletions does not caused significant change to the baseline estimate of 6.4%. The strongest downward effect had a deletion of Germany, which we explain by the fact that in this country most internal changes of the CEE policy took place. But, even in this case, the difference to the baseline estimate is not significant.

In the second set of our robustness analysis (Table 7), we experiment with alternative weighting schemes, setting the high-school graduation year, inclusion of additional dummies and different PIAAC sample treatment. The rows [1] and [2] of Table 7 repeat the benchmark specification with adjusted weights and without weights. The adjusted weights rescale the sampling weights in such a way that within each country the weights sum up to one, so that each country has the same contribution to the overall effect. The estimated CEE effect in row [1] is not statistically different from the baseline estimate. We argue that the country fixed effects cover sufficiently the fixed international variation. In the case of no weights (row [2]) the CEE estimate is slightly higher than the baseline, due to the fact that some countries with larger samples, e.g. Poland, USA, are mostly no CEE countries and have lower average country numeracy score. Our CEE dummy is based on the reported age when the last education degree was obtained. For individuals with last school degree being high school or less this was the age we used for calculating her school graduating cohort. For individuals having a tertiary degree, we constructed their high school graduation age as median age of the high school graduation in their country. To check for sensitivity of this assumption we performed two robustness checks. In the first one, we assumed that each individual graduated at mean graduation age of her country (row [3] of Table 7). In the [4]th row we assumed that all PIAAC individuals graduated from high school at age 18. The CEE effects hardly differ using these two alternatives. The next two rows of the robustness table (row [5] and [6], Table 7) include into the estimation the occupational (ISCO) and industry (ISIC) dummies. The CEE effects slightly decrease due to this extension. In the row [7] of Table 7 we ignore the within country variation in CEE introduction for Canada and Germany and introduce a dynamic CEE factor instead, so that within a country and graduation year, there is no variation in the CEE dummy. The CEE impact decreases due to this aggregation. Backed by this result we argue that if we had more detailed individual information on regions for the decentralized countries (Australia, Canada, USA), the baseline CEE impact would have been a little larger. Finally the [8]th robustness checks includes Russian PIAAC sample. The CEE effect increases due to the lower average numeracy scores in Russia and Russia being a CEE country only since 2009.

Table 8 repeats the baseline specification of Table 2, column (6) with different subgroups. Results for the different subsamples remain positive and mainly statistically significant which means that all individuals benefit from central exams in terms of cognitive skills. But nevertheless estimates differ substantially across subgroups and range from 4.2% to 10.8%. The estimates of CEE effect of males [1] exceeds the one of females [2] with 6.9% to 5.9 % of a standard deviation. This is in line with our baseline results where the estimate of female dummy features a negative sign. The CEE effect is somewhat higher for subgroups with weaker social backgrounds such as migration background and low educated parents. As second-generation migrants tend to have first-generation migrants as their parents, they are expected to have fewer opportunities in the acquisition of cognitive skills. According to our results CEEs seem to mitigate this inequality in educational opportunities as the CEE effect for second-generation migrants exceeds that of native individuals [4]. Also among the subgroups of parental education, the group of below high school parents [5] and high school parents [6] features the highest CEE effect compared to high educated parents [7]. This results are in line with the findings of Woessmann (2002) and support our baseline results.

6 Conclusion

This paper investigates the effect of centralized exit examinations (CEEs) at the end of secondary school on cognitive skills during adulthood using data from the recent OECD survey "Programme for International Assessment of Adult Competencies" (PIAAC

2013). While literature agrees that CEEs are associated with substantially higher test scores of students (Bishop 1997; Woessmann 2003, 2005; Jürges et al. 2005) critics warn that this result might simply reflect differences in students' test taking ability rather than actual differences in knowledge and skills (Popham 2001; Volante 2004).

In this study we merge the PIAAC adult skills data with the country or state specific information on CEE vs. local exam regimes at the end of the high school. Using the cross- and within-country variation in the test regimes, we estimate a stable positive effect of CEEs at the end of secondary education on cognitive skills of the adult population. Conditional on covariates, individuals in education systems with centralized exams at the end of high school are associated with more than 20 percent of a standard deviation higher numeracy skills during adulthood. Taking unobserved country variation into account, graduates of CEE regimes still significantly outperform graduates of local exam regimes by more than 6 percent of a standard deviation in numeracy and by more than 12 percent in literacy and problem solving scores. Our results are robust to a large set of alternative specifications and robustness checks. We also show that a CEE regime significantly increases the individual probability of employment and higher educational attainment. Exploiting variation in exam types across German federal states, we document CEE effects of similar magnitudes even within one country. Lastly, we estimate the CEE effects using data from the 2000 wave of the "Programme for International Student Assessment" (PISA) along the lines of the analysis by Fuchs and Woessmann (2007). Limiting the PIAAC data to the 2000 PISA cohort, we estimated CEE effects of roughly the same size in both datasets.

Altogether our results document an exemplary robust evidence of the persistent CEE effects on adult skills. Moreover the positive effects of CEE in employment and higher education attainment supports the hypothesis that CEEs are an effective educational device with positive consequences beyond the mere increase in the students' test results.

References

Backes-Gellner, U. and S. Veen (2008). The consequences of central examinations on educational quality standards and labour market outcomes. *Oxford Review of Education 34*(5), 569–588.

Becker, W. E. (1982). The educational process and student achievement given uncertainty in measurement. *American Economic Review 72*(1), 229–36.

Becker, W. E. and S. Rosen (1992). The learning effect of assessment and evaluation in high school. *Economics of Education Review 11*(2), 107–118.

Bishop, J. H. (1997). The effect of national standards and curriculum-based exams on achievement. *American Economic Review: Papers and Proceedings 87*(2), 260–264.

Bishop, J. H. (1998). The effect of curriculum-based external exit exam systems on student achievement. *The Journal of Economic Education 29*(2), 171–182.

Bishop, J. H. (1999). Are national exit examinations important for educational efficiency? *Swedish Economic Policy Review 6*, 349–398.

Bishop, J. H. (2006). Drinking from the fountain of knowledge: Student incentive to study and learn - externalities, information problems, and peer pressure. In E. A. Hanushek and F. Welch (Eds.), *Handbook of Labor Economics*, Handbook of the Economics of Education, pp. 909–944. Amsterdam: North-Holland.

Bishop, J. H. and L. Wößmann (2004). Institutional effects in a simple model of educational production. *Education Economics 12*(1), 17–38.

Boezerooy, P. and J. Huisman (2000). From secondary to tertiary education in the netherlands. *Journal of Institutional Research 9*(1), 54–64.

Caves, K. and S. Balestra (2014). The impact of high school exit exams on graduation rates and achievement. Working paper, UZH Business Working Paper Series, No. 346.

Fuchs, T. and L. Woessmann (2007). What accounts for international differences in student performance? a re-examination using pisa data. *Empirical Economics 32*, 433–464.

Greger, D. and E. Kifer (2012). Lost in translation: Ever changing and competing purposes for national examinations in the Czech republic. *Journal of Pedagogy / Pedagogický Časopis 3*(1), 43 –81.

Hanushek, E. A., G. Schwerdt, S. Wiederhold, and L. Woessmann (2015). Returns to Skills Around the World: Evidence from PIAAC. *European Economic Review 73*, 103–130.

Hörner, W., H. Döbert, B. von Kopp, and W. Mitter (2007). *The education systems of Europe*. Springer.

Jürges, H., K. Schneider, and F. Büchel (2005). The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany. *Journal of the European Economic Association 3*(5), 1134–1155.

Klein, E. D., S. M. Kühn, I. v. Ackeren, R. Block, et al. (2009). Wie zentral sind zentrale prüfungen? abschlussprüfungen am ende der sekundarstufe ii im nationalen und internationalen vergleich. *Zeitschrift für Pädagogik 55*(4), 596–621.

Lamprianou, I. (2012). Unintended consequences of forced policy-making in high stakes examinations: The case of the republic of cyprus. *Assessment in education: Principles, policy & practice 19*(1), 27–44.

Lüdemann, E. (2011). Intended and unintended short-run effects of the introduction of central exit exams. In E. Lüdemann (Ed.), *Schooling and the formation of cognitive and non-cognitive outcomes*, Ifo Beitrge zur Wirtschaftsforschung 39. Ifo Institut.

PIAAC (2013). Technical report of the survey of adult skills (PIAAC). Technical report, OECD. Available at www.oecd.org/site/piaac/_Technical%20Report_17OCT13.pdf.

Piopiunik, M., G. Schwerdt, and L. Woessmann (2013). Central school exit exams and labor-market outcomes. *European Journal of Political Economy 31*, 93–108.

Popham, W. J. (2001). Teaching to the test? *Educational Leadership 58*(6), 16–21.

Schwerdt, G. and L. Woessmann (2015). The information value of central school exams. CESifo Working Paper No. 5404, Munich: CESifo.

Volante, L. (2004). Teaching to the test: What every educator and policy-maker should know. *Canadian Journal of Educational Administration and Policy* (35).

Vraketta, G. (2013). A washback study of the pancyprian examinations on teacher stress. Mimeo, University of Leicester.

Woessmann, L. (2002). Central exams improve educational performance: International evidence. Kiel Discussion Papers 397, Kiel Institute for the World Economy (IfW).

Woessmann, L. (2003). Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics 65*(2), 117–170.

Woessmann, L. (2005). The effect heterogeneity of central examinations: evidence from TIMSS, TIMSS-Repeat and PISA. *Education Economics 13*(2), 143–169.

7 Figures and Tables

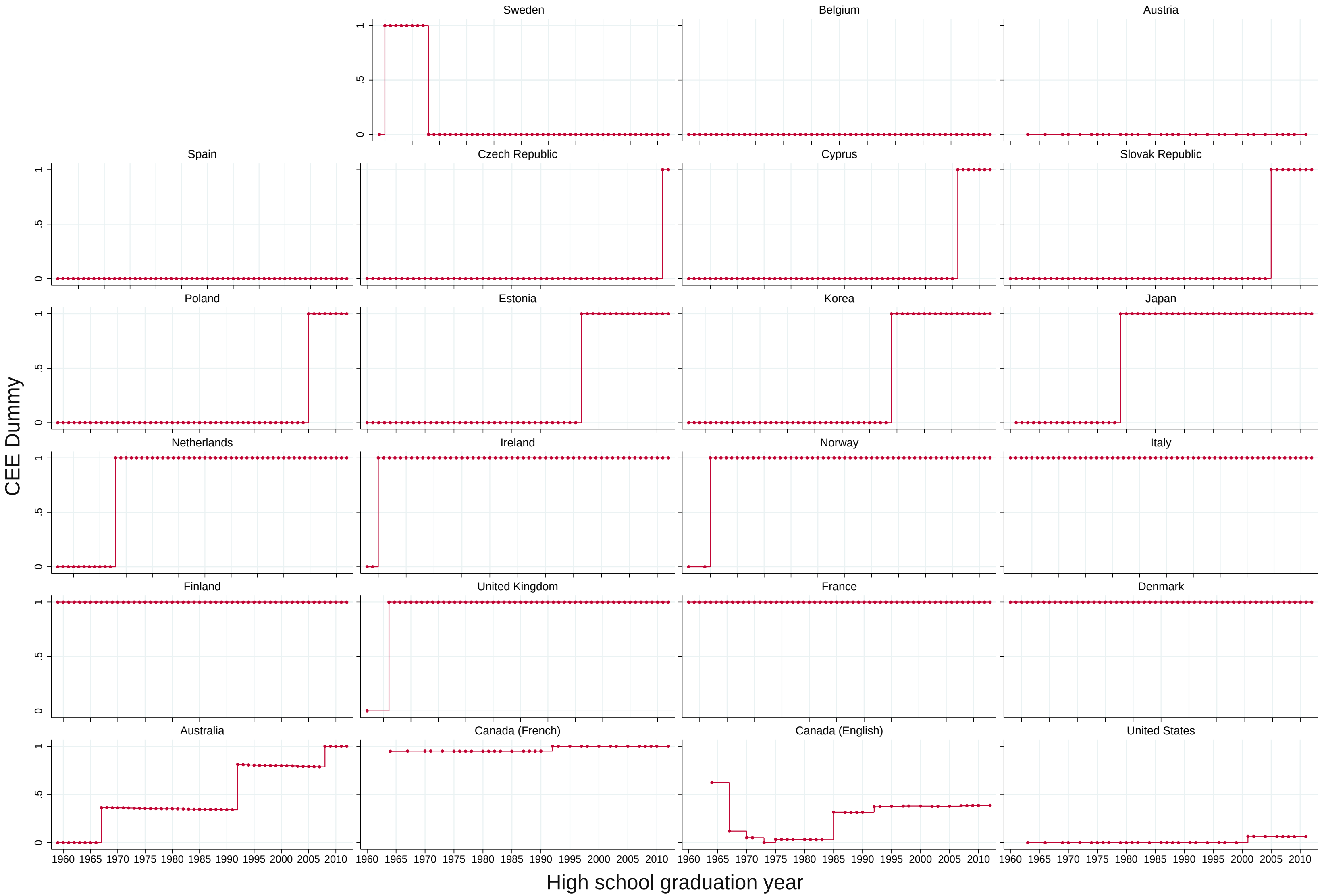

Figure 1: CEE regimes in PIAAC countries (1960-2012)

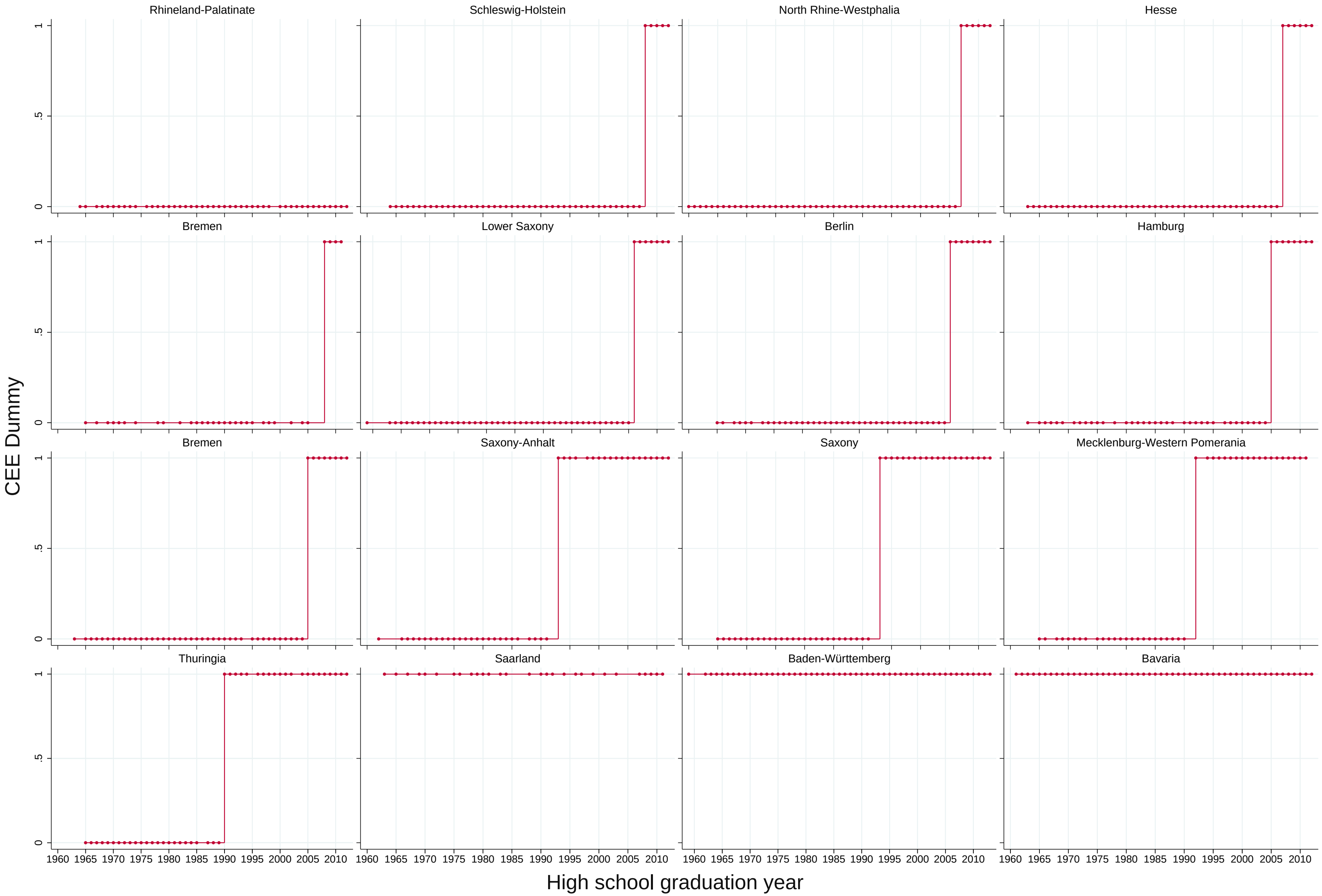

Figure 2: CEE regimes in German federal states (1960-2012)

Table 1: Descriptive statistics of the PIAAC sample

	Non-CEE		CEE			
Variable	Mean	*Std. dev.*	Mean	*Std. dev.*	Mean diff.	t-value
Nummeracy	-0.108	*0.0095*	0.11	*0.0059*	0.217	19.30
Literacy	-0.01	*0.0089*	0.162	*0.0058*	0.172	16.15
Problem-solving	-0.055	*0.0123*	0.244	*0.0083*	0.299	20.16
Earnings	2.8	*0.0074*	2.718	*0.0048*	-0.082	-9.27
Employment	0.903		0.917		0.014	3.31
ISCED 3+	0.851		0.81		-0.041	-10.08
ISCED 5+	0.342		0.345		0.004	0.68
Age	43.2	*0.1359*	36.899	*0.0820*	-6.302	-40.02
Age at graduation	18.526	*0.0329*	18.312	*0.0249*	-0.214	-5.19
Female	0.504		0.501		-0.004	-0.65
Second-gen. migrant	0.101		0.08		-0.021	-6.02
Parents educ.:						
Primary	0.265		0.3001		0.036	8.38
Secondary	0.406		0.371		-0.035	-6.30
Above secondary	0.284		0.25		-0.034	-6.41
Missing	0.046		0.079		0.033	12.88
Observations⋆	71,699		68,483			

Notes: The above means employ sampling weights. Primary sampling unit in PIAAC is an individual, while the strata is country. We standardized the three skills scores on the global level. For earnings we use log gross hourly wage, employment is a 0/1 variable, where 1 means currently employed. The two ISCED (International Standard Classification of Education) variables are educational dummies. ISCED 3+ equals 1 if an individual completed a high school, and ISCED 5+equals 1 if an individual completed any colleague.

⋆ The number of observations is smaller for following variables: for ISCEDs, because we exclude those currently studying; for earnings mean and employment share we additionally exclude those younger than 25; finally, the problem-solving questionnaire was not completed in 4 countries: Cyprus, France, Italy and Spain.

Table 2: Baseline estimates of the CEE effect

	(1)	(2)	(3)	(4)	(5)	(6)
CEE	.226***	.177***	.272***	.073**	.083***	.064***
	(.036)	*(.037)*	*(.035)*	*(.033)*	*(.027)*	*(.019)*
Age		.047***	.054***	.052***	.025***	
		(.010)	*(.009)*	*(.007)*	*(.008)*	
$\text{Age}^2/100$		–.067***	–.066***	–.066***	–.033***	
		(.011)	*(.010)*	*(.009)*	*(.009)*	
Female			–.228***	–.223***	–.220***	–.220***
			(.011)	*(.011)*	*(.011)*	*(.011)*
Second-gen migrant			–.028	–.017	–.020	–.013
			(.023)	*(.024)*	*(.022)*	*(.025)*
Parents educ.:						
Secondary			.436***	.415***	.322***	.310***
			(.019)	*(.024)*	*(.027)*	*(.026)*
Above secondary			.789***	.761***	.645***	.633***
			(.019)	*(.026)*	*(.033)*	*(.032)*
Educ. missing			–.160***	–.177***	–.173***	–.180***
			(.035)	*(.034)*	*(.033)*	*(.032)*
R^2	.012	.036	.147	.188	.234	.239
Fixed effects:						
country				yes	yes	yes
age at graduation					yes	yes
age						yes

Notes: Dependent variable is globally standardized nummeracy score. Nr. of observations is 140,182; nr. of countries is 23. Estimations are weighted by sampling weights. Robust standard errors (in parentheses) account for clusters made of graduation year × CEE region.
* $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$.

Table 3: Estimates of the CEE effect in other outcomes: specification (6) of Table 2

	(1) Literacy	(2) Prob. solv.	(3) Earnings	(4) Employed	(5) ISCED 3+	(6) ISCED 5+
CEE	.122***	.139***	.017	.024***	.008	.043***
	(.019)	*(.022)*	*(.015)*	*(.008)*	*(.005)*	*(.010)*
Female	–.016	–.120***	–.225***	–.012**	.009***	.018**
	(.011)	*(.015)*	*(.014)*	*(.005)*	*(.003)*	*(.008)*
Second-gen migrant	–.019	–.027	.051***	–.009	.002	.032***
	(.023)	*(.026)*	*(.015)*	*(.009)*	*(.007)*	*(.009)*
Parents educ.:						
Secondary	.302***	.344***	.131***	.017**	.070***	.117***
	(.019)	*(.029)*	*(.010)*	*(.008)*	*(.008)*	*(.010)*
Above secondary	.636***	.653***	.257***	.040***	.080***	.343***
	(.028)	*(.033)*	*(.018)*	*(.009)*	*(.008)*	*(.017)*
Educ. missing	–.142***	–.062	–.050***	–.056***	–.005	–.010
	(.026)	*(.041)*	*(.019)*	*(.015)*	*(.011)*	*(.010)*
R^2	.237	.176	.319	.053	.638	.319
Observations	140182	94113	61963	78298	92356	92356
Fixed effects:						
country	yes	yes	yes	yes	yes	yes
age at graduation	yes	yes	yes	yes	yes	yes
age	yes	yes	yes	yes	yes	yes

Notes: Dependent variables: for literacy and problem solving skills we use globally standardized scores. For earnings we use log gross hourly wage, employment is a 0/1 variable, where 1 means currently employed. The two ISCED (International Standard Classification of Education) variables are educational dummies. ISCED 3+ equals 1 if an individual completed a high school, and ISCED 5+equals 1 if an individual completed any colleague. The ISCED estimations excludes those currently studying, the earnings and employes estimation additionally excludes those younger than 25. Estimations are weighted by sampling weights. Robust standard errors (in parentheses) account for clusters made of graduation year × CEE region.
* $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$.

Table 4: Estimates of the CEE effect in different outcomes for German subsample

	(1) Numeracy	(2) Literacy	(3) Prob. solv.	(4) Earnings	(5) Employed	(6) ISCED3+	(7) ISCED5+
CEE	.112***	.099***	.057*	.008	.028***	.023***	.019
	(.028)	*(.030)*	*(.034)*	*(.023)*	*(.011)*	*(.008)*	*(.013)*
Female	−.248***	−.033	−.118***	−.218***	.009	.019***	−.065***
	(.028)	*(.030)*	*(.030)*	*(.021)*	*(.011)*	*(.007)*	*(.012)*
Second-gen migrant	.008	−.033	.002	.016	−.000	−.019*	.017
	(.034)	*(.037)*	*(.040)*	*(.028)*	*(.012)*	*(.010)*	*(.015)*
Eastern German	−.195***	−.112**	−.100*	−.252***	−.041*	.017	−.042*
	(.049)	*(.054)*	*(.053)*	*(.039)*	*(.024)*	*(.011)*	*(.023)*
Parents educ.:							
Secondary	.311***	.236***	.422***	.154***	.003	.050**	.041*
	(.059)	*(.058)*	*(.062)*	*(.044)*	*(.020)*	*(.020)*	*(.025)*
Above secondary	.562***	.541***	.649***	.235***	.006	.043**	.166***
	(.063)	*(.063)*	*(.067)*	*(.047)*	*(.020)*	*(.019)*	*(.027)*
Educ. missing	−.140	−.253***	.038	−.007	−.061*	−.043*	.021
	(.087)	*(.089)*	*(.096)*	*(.061)*	*(.034)*	*(.026)*	*(.030)*
R^2	.306	.287	.259	.277	.100	.590	.513
Observations	4137	4137	3669	2118	2436	3293	3293
Fixed effects:							
age	yes	yes	yes	yes	yes	yes	yes
age at graduation	yes	yes	yes	yes	yes	yes	yes

Notes: Dependent variables: for numeracy, literacy and problem solving skills we use globally standardized scores. For earnings we use log gross hourly wage, employment is a 0/1 variable, where 1 means currently employed. The two ISCED (International Standard Classification of Education) variables are educational dummies. ISCED 3+ equals 1 if an individual completed a high school, and ISCED 5+equals 1 if an individual completed any colleague. The ISCED estimations excludes those currently studying, the earnings and employes estimation additionally excludes those younger than 25. Graduates from the former GDR (Graduation year <1991) are excluded. Estimations are weighted by sampling weights. Robust standard errors (in parentheses) account for clusters made of graduation year × CEE region.
* $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$.

Table 5: Comparison estimates of the CEE effect in numeracy skills (PIAAC) and math scores (PISA/TIMSS)

	PIAAC					PISA		TIMSS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
CEE	.417***	.511***	.531***	.419***	.483***	.347**	.356**	.377*	.415*
	(.076)	*(.049)*	*(.049)*	*(.076)*	*(.074)*	*(.154)*	*(.160)*	*(.203)*	*(.231)*
Female	−.324***	−.332***	−.331***	−.326***	−.327***	−.096***	−.096***	−.072***	−.072***
	(.033)	*(.032)*	*(.033)*	*(.033)*	*(.032)*	*(.009)*	*(.009)*	*(.014)*	*(.014)*
Second-gen migrant	−.004	.045	.039	−.004	−.006	−.119**	−.119*	−.114**	−.113**
	(.088)	*(.057)*	*(.060)*	*(.089)*	*(.090)*	*(.056)*	*(.057)*	*(.047)*	*(.046)*
Eastern Europe	−.021	.458***	−.034	−.048	−.427***	−.118	−.464***	.245***	−.132
	(.089)	*(.060)*	*(.049)*	*(.077)*	*(.050)*	*(.082)*	*(.139)*	*(.024)*	*(.208)*
Parents educ.:									
Secondary	.552***	.612***	.611***	.551***	.554***	.244***	.242***	.424***	.423***
	(.127)	*(.086)*	*(.086)*	*(.129)*	*(.130)*	*(.058)*	*(.061)*	*(.085)*	*(.085)*
Above secondary	.985***	1.042***	1.041***	.984***	.987***	.641***	.636***	.780***	.780***
	(.157)	*(.124)*	*(.125)*	*(.159)*	*(.161)*	*(.065)*	*(.066)*	*(.097)*	*(.096)*
Educ. missing	.047	.065	.062	.046	.034	.626***	.622***	.079	.079
	(.165)	*(.151)*	*(.152)*	*(.167)*	*(.174)*	*(.173)*	*(.175)*	*(.188)*	*(.188)*
R^2	.154	.176	.175	.154	.156	.089	.089	.112	.111
Observations	4900	4291	4291	4573	4573	61029	61029	50758	50758
CEE coding	our	our	F& W	our	F&W	our	F&W	our	F&W
Nr. countries	23	13	13	20	20	20	20	13	13

Notes: The dependent variable is standardized numeracy score for PIAAC and standardized math score for PISA and TIMSS. PISA data are from the PISA 2000 wave. TIMSS data are from the 1999 TIMSS-Repeat wave. The estimations in PIAAC columns are based on variously restricted PIAAC sample to match it with the PISA and TIMSS datasets. Column (1), (4), and (5) are based on sample of PIAAC individuals who were 15 years old in 2000 (i.e. match to PISA), while columns (2) and (3) are based on sample of PIAAC individuals who were born in 1984-85 (i.e. match to TIMSS). The estimations in all columns are limited to the section of countries which are common to PIAAC/PISA and PIAAC/TIMSS. The columns (3), (5), and (7) impose the CEE coding adopted from the Fuchs and Woessmann (2007) (see A3), remaining columns take our CEE coding introduced in Section 2. Consequently the column (2) is comparable to (8), (3) to (9), (4) to (6), and (5) to (7). Column (1) takes all PIACC countries into estimations. All estimations are weighted by the respective sampling weights. Robust standard errors account for clustering on the country level. * $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$.

Table 6: Piecewise deletion of countries

Baseline	.064***	*(.019)*			
Excluded country:					
Australia	.067***	*(.019)*	Italy	.061***	*(.019)*
Austria	.065***	*(.019)*	Japan	.126***	*(.018)*
Belgium	.065***	*(.019)*	Korea	.041**	*(.021)*
Canada	.072***	*(.019)*	Netherlands	.066***	*(.019)*
Cyprus	.064***	*(.019)*	Norway	.063***	*(.019)*
Czech Republic	.066***	*(.019)*	Poland	.049**	*(.019)*
Denmark	.063***	*(.019)*	Slovak Republic	.063***	*(.019)*
Estonia	.064***	*(.019)*	Spain	.075***	*(.019)*
Finland	.065***	*(.019)*	Sweden	.063***	*(.019)*
France	.070***	*(.019)*	United Kingdom	.057***	*(.018)*
Germany	.041*	*(.022)*	United States	.066***	*(.019)*
Ireland	.064***	*(.019)*			

Notes: Dependent variable is globally standardized nummeracy score. CEE effects estimates and underlying standard errors (in parantheses) stemming from specifications (6) of Table 2 dropping one country. Robust standard errors (in parentheses) account for clusters made of graduation year × CEE region.
* $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$.

Table 7: Alternative specifications

Baseline	.064***	*(.019)*			
[1] Adjusted weights	.068***	*(.016)*	[5] ISCO dummies	.044***	*(.017)*
[2] No weights	.084***	*(.017)*	[6] ISIC dummies	.046***	*(.017)*
[3] Median age of graduation	.065***	*(.019)*	[7] Only countries	.050**	*(.024)*
[4] Age of graduation flat at 18	.062***	*(.019)*	[8] with Russia	.075***	*(.019)*

Notes: CEE estimates in alternative specifications of column (6) of Table 2. Dependent variable is globally standardized numeracy score. Robust standard errors (in parentheses) account for clusters made of graduation year × CEE region.
* $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$

Table 8: Subgroup analysis

Baseline	.064***	*(.019)*			
[1] Males	.069***	*(.024)*	*Parents educ.:*		
[2] Females	.059***	*(.021)*	[5] Below high school	.095***	*(.030)*
[3] Natives	.064***	*(.019)*	[6] High school	.119***	*(.024)*
[4] Second-gen migrants	.108**	*(.047)*	[7] Above high school	.042	*(.031)*

Notes: CEE estimates for various subgroups, we repeat here the specification (6) of Table 2. Dependent variable is globally standardized numeracy score. Robust standard errors (in parentheses) account for clusters made of graduation year × CEE region.
* $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$

Appendix

Table A1: Introduction of the Central Exam Examinations on the upper secondary level (ISCED 3) across PIAAC countries

Country	Year of CEE-Introduction	Source
Australia*	1992-2007	Australasian Curriculum, Assessment and Certification Authorities (ACACA): http://acaca.bos.nsw.edu.au/go/leaving-school/nsw/certificates
Austria	2016	Ministerium für Bildung und Frauen Österreich (BMBF): https://www.bmbf.gv.at/schulen/unterricht/ba/reifepruefung.html
Belgium	-	Hörner et al. (2007), Centre for Educational research, London School of Economics and Political Science: http://www.leeds.ac.uk/educol/documents/00001195.htm
Canada*	1929-1984	General Accounting Office, Washington, 1993: http://files.eric.ed.gov/fulltext/ED361377.pdf, https://archive.org/details/ERIC_ED361377
Cyprus	2006	Vraketta (2013), Lamprianou (2012), Hörner et al. (2007), Ministry of Education Cyprus: http://www.highereducation.ac.cy/en/examinations.html
Czech Republic	2011	email communication with Centre for Evaluation of Educational Achievement: www.cermat.cz, Greger and Kifer (2012)
Denmark	1908	Hörner et al. (2007), Bishop (1999)
Estonia	1997	Education fact sheets on http://archimedes.ee, http://www.innove.ee/en/general-education/final-examinations/secondary-school

*Countries with non-unified school systems. The CEE introduction years differ across federal states. See Table A2 for more detailed information.

Table A1: *Continued*

Country	Year of CEE-Introduction	Source
Finland	1919	Matriculation Examination Board Finland: https://www.ylioppilastutkinto.fi/fi/english https://en.wikipedia.org/wiki/Matriculation_exam_%28Finland%29
France	1808	Hörner et al. (2007), Ministry of Education France: http://www.education.gouv.fr/cid60987/bac-2015-questions-reponses.html
Germany*	1945-2007	Lüdemann (2011), Klein et al. (2009)
Ireland	1960	State examination Commission Ireland: https://www.examinations.ie/?l=en&mc=li&sc=li Department of Education and Skills: http://www.education.ie/en/The-Education-System/Post-Primary/
Italy	1923	personal communication
Japan	1979	personal communication
Korea	1994	personal communication
The Netherlands	1968	Hörner et al. (2007), Boezerooy and Huisman (2000), Government of the Netherlands: https://www.government.nl/topics/secondary-education/contents/secondary-school-leaving-examination
Norway	1960	Hörner et al. (2007), Organization for Internationalization in Education: https://www.epnuffic.nl/en/publications/find-a-publication/education-system-norway.pdf
Poland	2005	The Central Examination Commission: http://www.cke.edu.pl/, http://www.cke.edu.pl/images/stories/English/the_matura_exam.pdf
Spain	-	Bishop (1999), Fuchs and Woessmann (2007), Hörner et al. (2007)
Slovak Republik	2005	National Institute for Certified Educational Measurement: http://www.nucem.sk/en/maturita http://www.nucem.sk/en/maturita#3/brief-history-of-maturita
Sweden	eliminated in 1968	Bishop (1999), Hörner et al. (2007), https://en.wikipedia.org/wiki/Studentexamen
United Kingdom	in 1950's	https://en.wikipedia.org/wiki/Higher_School_Certificate_(United_Kingdom) http://www.a-levels.co.uk/history-of-a-levels.html
United States*	2001	Caves and Balestra (2014)

*Countries with non-unified school systems. The CEE introduction years differ across federal states. See Table A2 for more detailed information.

Table A2: Central Exam Examinations on the upper secondary level (ISCED 3) in countries with non-unified school systems

Country	Federal state	Year of CEE-Introduction
Australia	New South Wales	1967
	Northern Territory	1992
	Queensland	2008
	South Australia	1992
	Tasmania	1992
	Victoria	1992
	Western Australia	1992
Canada	Alberta	1984
	British Columbia	1984
	Manitoba	eliminated 1970/reintroduced 1991
	Newfoundland/Labrador	1974
	New Brunswick (francophone)	1991
	New Brunswick (anglophone)	-
	Northwest Territories	-
	Nova Scotia	eliminated 1972
	Nunavut	-
	Ontario	eliminated 1967
	Prince Edward Island	-
	Quebec	1929
	Saskatchewan	-
	Yukon	-
Germany	Baden-Württemberg (BW)	1952
	Bavaria (BY)	1946
	Berlin (BE)	2006
	Brandenburg (BB)	2005
	Bremen (HB)	2007
	Hamburg (HH)	2005
	Hesse (HE)	2007
	Mecklenburg-Western Pommerania (MV)	1991
	Lower Saxony (LS)	2006
	North-Rhine Westphalia (NW)	2007
	Rhineland Palantine (RP)	-
	Saarland (SL)	1945
	Saxony (SN)	1993
	Saxony-Anhalt (ST)	1993
	Schleswig-Holstein (SH)	2008
	Thuringia (TH)	1990
USA*	New York	2002

*Within the USA, only in the federal state of New York has a centralized school leaving exam that meets our definition.

Table A3: Fuchs and Woessmann (2007) and our coding of the CEE factors at the end of secondary school across common PIAAC/PISA countries in year 2000

Year	F&W	Our
Australia	0.8	0.798
Austria	0	0
Belgium	0	0
Canada	0.5	0.537
Czech Republic	1	0
Denmark	1	1
Finland	1	1
France	1	1
Germany	0.4	0.424
Ireland	1	1
Italy	1	1
Japan	1	1
Korea	1	1
Netherlands	1	1
Norway	1	1
Poland	1	0
Spain	0	0
Sweden	0.5	0
United Kingdom	1	1
United States	0.1	0.067