University of Konstanz
Department of Economics

The Information Value of Central School Exams

Guido Schwerdt and Ludger Woessmann

Working Paper Series
2015-14

http://www.wiwi.uni-konstanz.de/econdoc/working-paper-series/

The Information Value of Central School Exams*

Guido Schwerdt† Ludger Woessmann‡

Abstract

The central vs. local nature of high-school exit exam systems can have important repercussions on the labor market. By increasing the informational content of grades, central exams may improve the sorting of students by productivity. To test this, we exploit the unique German setting where students from states with and without central exams work on the same labor market. Our difference-in-difference model estimates whether the earnings difference between individuals with high and low grades differs between central and local exams. We find that the earnings premium for a one standard-deviation increase in high-school grades is indeed 6 percent when obtained on central exams but less than 2 percent when obtained on local exams. Choices of higher-education programs and of occupations do not appear major channels of this result.

Keywords: Central exit exams, labor-market sorting, earnings, measurement error, difference-in-difference, Germany

JEL classification: I20, J24, J31

June 8, 2015

---

* This research has profited from comments by John Bishop, Sue Dynarski, David Figlio, Eric Hanushek, Joop Hartog, Steve Rivkin, Kerstin Schneider, Marty West, and Thomas Zwick, as well as seminar participants at Stanford, ETH Zurich, Leipzig, the Winter Meeting of the North American Econometric Society, the annual meetings of the EEA, EALE, ESPE, SMYE, the German Economic Association, and its economics of education committee, the CESifo area conference on economics of education, the IFN Vaxholm schooling conference, the economics of education workshop at the University of Duisburg-Essen, the central exams workshop at the Ifo Institute in Munich, and the educational governance workshop at the NTNU in Trondheim. Financial support by the German Federal Ministry of Education and Research for the project "Central exit exams as a governance instrument in the school system: Labor market effects, transmission channels, and effect heterogeneity" within the BMBF research priority "Educational governance" (SteBis) is gratefully acknowledged.

† Department of Economics, University of Konstanz, Box D133, 78457 Konstanz, Germany; also CESifo and IZA; guido.schwerdt@uni-konstanz.de.

‡ University of Munich, Ifo Institute, Poschingerstr. 5, 81679 Munich, Germany; also CESifo, IZA, and CAGE; woessmann@ifo.de.

1. Introduction

A key question of educational governance is how to organize school exit examinations. Historically exams were designed and graded by local school teachers, but today centralized exams at the end of secondary school play an important role in many education systems around the world. Supporters argue that centralized evaluations are desirable because they raise student achievement by favorably affecting the incentive structure of stakeholders in the education system. Empirical evidence largely supports such a positive effect of central exams on student achievement.[1] But externally validated credentials may also have important repercussions on the labor market. By increasing the informational content of grades, central exams could affect school-to-work transitions, college admissions, and hiring costs and hiring policies of firms – and, by implication, the incentive structure in schools in the long run. However, whether grades obtained in central exams indeed allow an improved sorting of students by productivity remains an open question.

This paper tests empirically whether grades in central exams have higher information value on the labor market than grades in local exams. The key empirical challenge consists in estimating the counterfactual relationship between high-school grades and productivity. Wages provide a measure for productivity that is readily available in many surveys of the labor force, but information on high-school grades is typically not. More importantly, the extent to which wages proxy for productivity in different countries may differ due to different labor market institutions, making cross-country comparisons difficult to interpret. However, any within-country approach typically lacks variation in the centrality of exam systems.

In this paper, we exploit the institutional setup in Germany in the 1990s for identification. At that time, almost half of the German states tested students in state-wide administered central exams at the end of high school, while the other half did not. While individual states in Germany are largely responsible for their education systems, labor market institutions are largely determined at the federal level. Thus, we can compare individuals who compete on the same labor market, but part of whom obtained their high-school grades in states with central exams and part of whom in states without central exams.[2]

[1] E.g., Bishop (1997, 2006), Woessmann (2003, 2005), Jürges, Schneider, and Büchel (2005), Jürges and Schneider (2010), and Jürges et al. (2012); see Hanushek and Woessmann (2011) for a review of the extensive cross-country evidence.

[2] Note that this is a different kind of variation from comparing U.S. states with and without minimum competency exams or differing graduation requirements (e.g., Bishop and Mane (2001); Dee and Jacob (2007); Baker and Lang (2013)), as all German states have such graduation requirements. Here, we are able to hold the existence of graduation requirements constant in order to test whether the external character of examinations has particular information value.

We use the dataset of a German university graduate survey whose participants graduated from university in the late 1990s. We observe their labor-market outcomes five years after graduation, in 2003. By surveying their high-school grades and following them into the labor market, this unique dataset for the first time allows us to convincingly estimate whether grades obtained in central exams indeed allow an improved sorting of students by productivity.

We test this hypothesis in a difference-in-difference framework. Identification is based on comparing earnings of individuals with high and low high-school grades depending on whether they obtained their grades in central or local exam states. Based on cross-sectional models, we estimate an earnings function with state fixed effects and focus on the interaction of an indicator for central exam states with standardized high-school grades. The hypothesis is that high-school grades are more closely associated with earnings when obtained in a central rather than local exam.

We find that a grade improvement by one standard deviation translates into approximately 6 percent higher earnings when grades are obtained in central exams, compared to 1.6 percent when grades are obtained in local exams. We interpret this finding in light of a simple measurement-error model in which grades proxy for productivity with classical measurement error. If local exam grades provide a more noisy measure of productivity compared to central exam grades, the resulting difference of more than 4 percentage points reflects the higher information value of central school exams.

This finding is robust to different sample selection criteria and alternative model specifications. Results hold for different subgroups of the overall sample. In line with economic intuition, the relationship between central exam grades and earnings is largest when we restrict the sample to individuals employed in the private sector. Controlling for several post-secondary schooling variables suggests that choices of higher-education programs and of occupations are no major channels driving our results.

The theoretical foundation for our analysis is laid in the literature on signaling educational performance, which builds on the seminal contributions by Spence (1973), Stiglitz (1975), and Arrow (1973). In models such as Bishop and Woessmann (2004) and Bishop (2006), central examinations change students' incentive structure relative to autonomous local examinations. By creating comparability to an external standard, central examinations improve the signaling of academic performance to advanced educational institutions and to potential employers. These institutions will thus give greater weight to educational performance when making admissions and hiring decisions. In consequence, their decisions become less sensitive to other factors such as family connections, the momentum of a twenty-minute job interview, performance relative to a class mean, or aptitude tests that measure innate ability more than overall educational performance. As students' rewards for learning grow, students respond by increasing their learning efforts. The idea that central exams increase the extrinsic rewards for learning is, however, an assumption in these models that has not yet been tested. It is one key contribution of this paper to provide the first empirical test of this assumption.

Our paper also directly contributes to the empirical literature on estimating labor-market returns to cognitive skills (see Hanushek and Rivkin (2012) for a summary). For Germany, Hanushek et al. (2015) report estimates of 24 percent for the earnings effect of a one standard deviation increase in test scores. Compared to this finding, our estimates of the relationship between earnings and school grades are small. The difference may arise from the restriction of our sample to successful university graduates, but it may also be an indication that school grades measure productive skills with substantially more noise than results of direct tests of the cognitive skills of the adult population.

Finally, our findings inform educational policy makers about repercussions on the labor market that a centralization of evaluation systems in schools may entail. Not only may central exam systems act as an effective accountability devise raising student achievement (see Hanushek and Woessmann (2011) for an overview of the international evidence) and overall labor-market productivity (Piopiunik, Schwerdt, and Woessmann (2013)). They also allow an improved sorting of students by productive skills. In the long run, this may increase overall welfare also in other ways. For example, central exams may facilitate the matching process between vacancies and workers (e.g., Petrongolo and Pissarides (2001)). Additional descriptive evidence on the relationship between the type of evaluation system in schools and applications, interviews, and job offers during the initial job search supports this conjecture. Central exams may also lower firms' costs of filling a vacancy because firms are less dependent on their own costly screening devises such as assessment centers. Finally, grades on school leaving exams are officially used to assign places at German universities when there is oversubscription in a subject or faculty. Thus, the matching process between scarce places in higher education programs and students with adequate skills may become more efficient. This may not only raise overall welfare, but also fairness in the access mechanism to the higher education system.

In what follows, section 2 provides a simple conceptual framework for the information value of central exam grades on the labor market. Section 3 describes the examination systems at the end of high school in German states, introduces the dataset, describes the sample, and provides descriptive statistics. Section 4 presents our empirical strategy. Sections 5 and 6 report our basic results and robustness tests. Section 7 tests to what extent university and job characteristics are channels of the effect. Section 8 presents supporting evidence on the higher information value of central exams based on the number of applications, interviews, and job offers during the initial job search. Section 9 concludes.

2. Conceptual Framework on the Information Value of Exams

2.1 Theoretical Background

Our analysis builds on the theoretical literature on signaling educational performance. The idea of job market signaling was introduced in the seminal work by Spence (1973): High-ability workers differentiate themselves from observationally identical workers of lower ability by acquiring an educational signal that is observed by potential employers. If higher ability individuals find it less costly (both in monetary and non-monetary terms) to acquire an educational degree, this process will lead to a separating equilibrium in which workers can be differentiated by their signals. The simple signaling model can be regarded as a special case of the more general screening model developed by Stiglitz (1975). In an economy with imperfect information, heterogeneous jobs demanding more or less skilled labor and workers with heterogeneous skill levels, education can enhance allocative efficiency by serving as a screening device that improves the job-worker matching. Arrow (1973) presents a similar idea in the filter theory of higher education where college education serves as a screening device that sorts out individuals of differing abilities and thereby conveys information to potential employers.

A more formal incorporation of the signaling and screening idea into models of educational production is presented in Becker (1982). His model extends the screening model of Stiglitz (1975) by explicitly linking post-school income to colleges' assessments of a student's academic achievement and prospective employers' assessments of the college the student attends. In particular, both the accuracy of student grading by the college and the accuracy of labeling of colleges by potential employers are functions of the screening process. Becker and Rosen (1992) investigate more explicitly the implications of student assessment schemes based on competition among peers on the one hand and an externally set competency standard on the other. The framing of education in a principal-agent setup emphasizes the role of incentives in the process of educational production. This setup also underlies studies that more directly focus on the relationship between educational standards and earnings. Costrell (1994) and Betts (1998) analyze the optimal setting of educational standards when students from schools with different grading standards are pooled. They conclude that centralized standard setting with a local option to set even higher standards results in higher standards, higher achievement, and higher social welfare than decentralized standard setting.

In a framework closest to our analysis, Bishop and Woessmann (2004) model how central examinations change students' incentive structures relative to autonomous local examinations. By creating comparability to an external standard, central examinations improve the signaling of academic performance to advanced educational institutions and to potential employers. These institutions will thus give greater weight to academic performance when making admission and hiring decisions. In consequence, their decisions become less sensitive to other factors such as family connections, short job interviews, performance relative to the class mean, or aptitude tests that lean more to measuring innate ability than to measuring overall educational performance. This is evident, for example, in Japan, Singapore, and South Korea, where performance on central high-school examinations directly determines whether or not students can proceed to tertiary education. Hence, central exams have a positive effect on the rewards for learning, especially on the extrinsic part. As students' rewards for learning increase, anything that increases educational performance becomes more worthwhile. Students respond to an increase in rewards by increasing their learning effort, and governments respond by increasing educational spending. The result is an increase in educational performance. In this model framework, a crucial assumption is that central exams have higher information value. This assumption is tested in this paper.

2.2 Central Exams in a Measurement-Error Framework

The key hypothesis tested in this paper is whether grades obtained in central exams have a higher information value which allows an improved sorting of students by productivity. Empirically, we test this hypothesis in a difference-in-difference framework. To provide a precise interpretation for our difference-in-difference parameter and to define the information value of central exams, we present a simple framework based on a model with classical measurement error in the explanatory variable.

Central exams are hypothesized to improve the reliability of the information on effective human capital to employers. We capture this idea by assuming that high-school grades can be thought of as a noisy signal of a worker's true productive skills, but grades obtained from central exit exams carry less noise than grades obtained from non-central exams. Accordingly, we define a grade obtained from a central exam, g_i^c, as a linear combination of productive skills a_i and a common noise component, n_i:

$$g_i^c = a_i + n_i \qquad (1)$$

We model the additional noise component introduced by non-central exams, e_i, as white noise. This can be interpreted as a classical measurement error problem. Grades obtained in a non-central exit exam, g_i^{nc}, can therefore be expressed as:

$$g_i^{nc} = a_i + n_i + e_i \qquad (2)$$

This setup allows us to define the information value of central school exams within a measurement-error context. For simplicity and without loss of generality, we assume the common noise component, n_i, to be zero. Abstracting also from other covariates, a regression of wages on ability is given by:

$$w_i = \beta a_i + \varepsilon_i \qquad (3)$$

Assuming that the measurement error, e_i, and the equation error ε_i are uncorrelated, standard classical measurement-error theory (e.g., Angrist and Krueger (1999)) implies that a regression of wages on high-school grades for graduates with non-central exit examinations can be expressed as:

$$w_i = \beta\lambda g_i^{nc} + \tilde{\varepsilon}_i \qquad (4)$$

where $\lambda = Cov(g^{nc}, a) / Var(g^{nc})$. If high-school grades for graduates with non-central exit exams proxy for productive skills with classical measurement error, then $Cov(g^{nc}, a) = Var(a)$ and $Var(g^{nc}) = Var(a) + Var(e)$, so the regression coefficient is necessarily attenuated, with the proportional "attenuation bias" equal to $(1 - \lambda) < 1$.

Hence, the difference in the coefficients from regressions of wages on high-school grades for graduates from central as opposed to non-central exit examinations would be given by $(1 - \lambda)$. The information value of a central exit examination could then be interpreted as a reduction of the "attenuation bias" in a classical measurement model.

The model builds on the common assumption that differences in wages reflect, at least to some extent, differences in productivity. This assumption requires that employers can observe on-the-job productivity. While this seems less plausible for starting wages, Altonji and Pierret

(2001) show that with the passage of time, employers learn about the true productive skills of workers. Thus, our analysis will mainly focus on earnings differences observed five years after university graduation.

3. Examination Systems in Germany and Graduate Data

3.1 Examination Systems at the End of High School in German States

Young adults in Germany take final exams in different subjects at the end of their secondary education, usually after 12 or 13 years of schooling. The official term in Germany for this certificate obtained at the end of the highest track of secondary education is *Allgemeine Hochschulreife*, commonly also labeled *Abitur*.[3] The certificate, issued after candidates have passed their final exams, enables individuals to attend university. Other school-leaving certificates from lower tracks of secondary education do not allow their holders to matriculate at a university.[4] In this regard, the *Abitur* serves the purpose of being the high-school leaving certificate as well as being a university entrance exam.

The final grade of the *Abitur* is a weighted average of the grades obtained in the final exams and grades obtained in courses taken during the two years before graduation.[5] The composition of the courses taken and the subjects of the final exams depend on choices of the individuals that are, however, restricted by certain regulations. The regulative framework for the choice of subjects varies between German states.

The particular institutional feature that we exploit in our empirical strategy is that the examination procedure for the final exams also varies between federal states in Germany. While in some federal states these final exams are external exit examinations, other states place the responsibility for the examinations entirely in the hands of the schools. In other words, in these states the candidates are examined by their respective teachers. The teachers formulate the examination questions and also grade the given answers. While nowadays, most German federal states have introduced external exit examinations, more than half of all federal states still had local exit examinations until 2005.

Table 1 provides official statistics on the *Abitur* in the German states distinguished by the type of exit examination by federal state. There are seven states with central exit exams and

[3] *Allgemeine Hochschulreife* could be translated as "general maturity for university studies". The term *Abitur* emanates from the Latin verb *abire* (to go away).

[4] Individuals holding certificates of *Hauptschulabschluss* or *Realschulabschluss* have the possibility to obtain a specialized *Fachabitur* ("maturity for specific university studies") or the *Abitur* if they graduate from a *Berufsschule* and then additionally attend a *Berufsoberschule*.

[5] Grant (2007) shows that averages of several grades provide useful information on the productivity of students.

nine states without them. This variation in institutional environments constitutes the main ingredient of our identification strategy. The table reveals significant variation in cohort shares and pass rates across states as well as variation between the two groups of states defined by the state-specific exit examination. In states with central exams, fewer students attend the highest track of secondary education, but those who do have slightly higher pass rates compared to students in the highest track in local-exam states.

3.2 Graduate Survey Data

Selected cohorts of university graduates have been included in large representative surveys conducted by the Hochschul-Informations-System (HIS, now called German Centre for Research on Higher Education and Science Studies or DZHW). Results were published in several descriptive studies (e.g., Kerst and Minks (2005)). In this paper, we use data for the 1997 cohort. University graduates of 1997 were surveyed in a first wave about one year after graduation (1998) and again in a second wave about five years after graduation (December 2002 - May 2003).[6]

The universe of individuals graduating from German universities in the academic year 1997 (September 1996 - September 1997) encompasses 191,948 individuals. Out of this universe, 9,583 graduates were initially sampled in 1998 and 6,220 individuals were sampled again in the second wave.[7] Our dataset contains information on the 6,216 graduates who were sampled in both waves.[8] This corresponds roughly to 3.2 percent of all graduates in 1997. We restrict the sample further by excluding those graduates who obtained their university access authorization outside Germany. Moreover, we exclude those who completed their secondary education in the former GDR. These restrictions are necessary to ensure the comparability of high-school grades. Altogether 359 observations were dropped due to these sample restrictions.

The dataset contains detailed information on personal characteristics, information on the course of university studies, on job search, and labor-market performance. Most importantly for our purpose, participants in the survey also reported their high-school grades (*Abitur* grades)[9] and the federal state in which they obtained their *Abitur*. This enables us to identify whether the final exams were external or local examinations. Moreover, we can link this

[6] The data are available as a scientific use file (ZA 4272) by the Gesis-ZA Zentralarchiv für Empirische Wirtschaftsforschung.

[7] Sampling follows a clustered sampling strategy. Clusters are defined by university, field of study, and type of degree.

[8] Four observations had to be dropped due to anonymity concerns.

[9] Henceforth, for simplicity we will refer to high-school grades instead of *Abitur* grades.

information on high-school grades and examination type to labor-market outcomes five years after graduation from university.

3.3 Descriptive Statistics

Table 2 visualizes mean differences in relevant characteristics between groups of individuals defined by type of examination and the relative position in the grade distribution. For illustrative purposes, the table distinguishes only between individuals with "good" and "bad" grades, defined by being above or below the median grade. Entries in the table represent mean values and standard deviations (in parentheses).

Comparing mean values of age, the share of males, and the share of individuals graduating in West-German states reveals only small differences between groups. Slightly larger differences exist in the share of graduates from a *Gymnasium* and graduates with *Fachabitur*. The share of graduates holding a *Fachabitur* amounts to only 13 percent in the non-central exam group compared to 19 percent among central exam graduates. Moreover, while the difference in shares of *Fachabitur* holders between graduates with bad and good grades is small in the non-central exam group, the respective difference is 11 percentage points among central exam graduates. Furthermore, there is a notable difference of 12 percentage points in the share of graduates from a *Gymnasium* between good and bad graders in the central exam group. These differences emphasize the importance of controlling for compositional differences along the *Fachabitur* and *Gymnasium* dimensions.

Table 2 also provides information on differences in average grades and earnings between groups. We normalized grades to have a mean of zero and a standard deviation of one. The average grade among graduates from central exam is 23 percent of a standard deviation above the mean, while the average grade in the local exam group is 9 percent of a standard deviation below. More importantly, mean differences between the groups of individuals with good and bad grades are almost identical. As variation in grades is the second key ingredient of our identification strategy, Figure 1 provides additional evidence on the distribution of demeaned grades by examination system. The figure suggests that the distribution of grades is roughly identical in both groups defined by the type of examination.

Table 2 further shows that central exam graduates also have higher earnings five years after graduation on average compared to local exam graduates. The difference in monthly earnings is roughly 90 Euros. Within examination type, earnings also vary substantially by exams grades, but the reward for being in the group of individuals with good grades is three times as high in states with central external examinations as it is in states with local exams.[10]

Figure 2 previews our difference-in-difference results by plotting average monthly earnings against grade categories separately for central exam and non-central exam states. The figure provides eye-ball evidence for a difference in the correlation between high-school grades and earnings by type of examination. While only a slight and somewhat noisy positive relationship between better grades and earnings is apparent among graduates from non-central exam states, the pecuniary reward from obtaining a good grade under a central exit examination regime is high.

Analyzing wage differentials between groups requires taking into account potential differences in labor supply as selection into employment might partially drive results (Heckman (1979)). At the bottom of Table 2, we therefore also report group-specific employment rates. Observed differences in employment rates as well as hours worked are, however, small and give little reason to be concerned about selection issues. We nevertheless deal with this concern more rigorously in the robustness section.

4. Empirical Model

The key challenge to studying whether grades obtained in central exams allow an improved sorting of students by productivity consists of estimating the counterfactual relationship between high-school grades and productivity. The unique institutional setup in Germany allows us to obtain an estimate of the counterfactual that is based on the observed relationship between grades and earnings within a control group of workers who obtained their high-school degree in German states with local exams. Thus, workers in this control group compete on the same labor market. By exploiting this setting, we can estimate the differential effect of high-school grades on earnings by type of examination based on a difference-in-difference framework within one country.[11]

The empirical model is given as:

$$y_i = \alpha + \beta \, Grade_i + \gamma \, CenExam_i + \delta \left(Grade_i * CenExam_i \right) + X_i ' \theta + \varepsilon_i \qquad (5)$$

[10] Figure A-1 in the appendix also reveals that the distributions of demeaned earnings within the two groups are very similar.

[11] Our framework differs from a generic difference-in-difference setup with one difference varying at the level of the observation. Moreover, we include grades in the multi-valued form instead of creating a binary variable. However, results reported in the appendix show that all our results hold when estimating a more traditional difference-in-difference approach.

where y_i represents log monthly earnings five years after graduation from university, $Grade_i$ is the high-school grade, $CenExam_i$ is an indicator variable for grades obtained in central exams, X_i is a vector of controls for personal characteristics, and ε_i is an error term uncorrelated with all right-hand-side variables. In equation (5), β reflects the effect of high-school grades on earnings in states with local exit examinations. The parameter γ captures the isolated effect of curriculum-based external exit examinations. In most of our empirical specifications, this effect will be subsumed by the inclusion of a set of fixed effects for federal states.

The key parameter of interest, δ, is identified by the estimate for the interaction effect between $Grade_i$ and $CenExam_i$. The parameter captures any additional effect of the high-school grade on earnings when grades are obtained in an external rather than local examination. In light of the measurement-error model presented above, the coefficient estimate for the parameter δ can be interpreted as the reduction in the "attenuation bias" due to the less noisy measurement of productivity-relevant skills.

Our identification strategy rests on the comparison between graduates from central exam states and graduates from non-central exam states. More precisely, the identifying assumption is that, conditional on other observed factors, the effect of high-school grades on earnings for graduates from non-central-exam states can serve as a valid estimate for the counterfactual grade-on-earnings effect for graduates from central-exam states in case they had not obtained their degree in a school system with external exit examinations. While this assumption cannot be tested, it should be noted that state-specific differences in earnings levels are captured by the inclusion of state fixed effects. Differences in regional labor markets or other institutional factors between states potentially bias our results only if they have a differential effect on earnings by the level of the high-school grade obtained.

5. Main Results on the Information Value of Central School Exams

The main results of our baseline model are shown in Table 3. Across the different specifications, there is clear evidence that better high-school grades are positively related to earnings and that this relationship is substantially stronger when these grades were obtained in central rather than local exams. The first column shows the basic difference-in-difference specification. The subsequent three specifications include fixed effects for the state of the high school. This set of dummy variables captures all grade-invariant variation between states. In the presence of mobility costs, comparison of labor-market outcomes of high-school graduates from states with and without central exams might be biased due to differences in local labor-market conditions. The specifications in the last two columns control also for other confounding factors. The third column introduces controls for family background and the fourth column adds the type of high-school degree and the degree-issuing institution to the list of control variables. All estimations allow for clustered standard errors by federal state.[12]

The main effects of interest are presented in the first two rows of the table. The estimate for the baseline effect of grades on earnings is statistically significant at 0.015-0.017 in all four columns. It suggests that when school grades are obtained in local exams, a grade improvement by one standard deviation is associated with 1.5-1.7 percent higher earnings on average. The interaction effect shows that this relationship is significantly different in states with central exams, where the association between high-school grades and future earnings is much stronger. The estimates for the interaction effect range from 0.034 to 0.043. They suggest that a grade improvement by one standard deviation translates into 6 percent (column 4) higher earnings in central-exam states.

In light of the model presented above, the estimated interaction effect can be technically interpreted as the reduction in the "attenuation bias" in the estimated coefficient of high-school grades due to the reduction in the noise component of grades obtained in central exit exams. Thus, the interaction effect reflects the additional information value of central school exams as opposed to local exams. The fact that its coefficient estimate is positive and significant supports our key hypothesis that grades obtained in central exams allow an improved sorting of students by productivity.

Before investigating the robustness of this finding, we quickly discuss the other parts of the empirical model. The estimated coefficients of the other control variables have the expected signs. Our results confirm that males earn more than females, which is in line with previous evidence on the existence of a gender wage gap in Germany (e.g., Fitzenberger and Wunderlich (2002); Reimer and Schröder (2006)). The significant negative effect of age seems striking at first glance. Most empirical studies investigating the age-earnings profile find a concave pattern suggesting a positive age effect for early and mid-career workers (e.g., Daveri and Maliranta (2007); Haltiwanger, Lane, and Spletzer (1999)). However, note that by focusing solely on persons graduating from university in the same year, we implicitly introduce a restriction for potential work experience after graduation, and those graduating at an older age might have needed more time to obtain their degree.

[12] When the number of groups is small, correcting the standard errors for within-group correlation might be a cure that is worse than the disease in a difference-in-difference estimation (Donald and Lang (2007)). However, all our results hold also without clustering.

Finally, in the first column we include an indicator variable for states with central examination systems instead of state fixed effects. The insignificant coefficient on the *CenExam* dummy reveals that central-exam states are not associated with higher earnings on average after controlling for differences in high-school grades, age, and gender composition in our sample. This is in line with existing evidence that suggests that for *Abitur* holders, there are no systematic earnings differences between individuals who obtained their high-school degree in states with central or local exams (see Backes-Gellner and Veen (2009); Piopiunik, Schwerdt, and Woessmann (2013)).

6. Sensitivity Analyses

This section investigates the robustness of our results. A first concern might be that the effect of high-school grades on earnings is heterogeneous and that our results might be driven by large effects for specific subgroups. We address this concern by replicating the above analysis for different subgroups of the population.

Table 4 provides regression results of our main model for seven different subgroups. All results are based on the most elaborate specification of equation (5), which corresponds to column 4 of Table 3. The first column of Table 4 shows estimates for individuals from the former West German states only. The results remain qualitatively unchanged with a slightly lower point estimate for the interaction effect of grades and the *CenExam* dummy.

Columns 2-5 present results for the "most common" scenario for obtaining a university-access degree according to the type of degree, the age at which the degree was obtained, and the degree-issuing institution. The most common scenario would imply obtaining the general university-access degree (*Abitur*) at an age between 18 and 20 at a *Gymnasium*. Hence, column 2 restricts the estimation sample to holders of a general university-access degree (*Abitur*), column 3 to individuals who obtained their university access degree at an age between 18 and 20, column 4 to individuals who obtained their degree at a *Gymnasium*, and column 5 imposes these three sample restrictions simultaneously. In all specifications, the estimate of the information value of central school exams stays rather constant.

Column 6 presents results based on an estimation sample for individuals working fulltime in the year of observation. Differences in the selection into fulltime employment by grade level and federal state may significantly affect the estimates of our baseline specification. However, column 6 reveals that the size of the estimate of the information value of central school exams declines by almost half once we restrict the sample to individuals working fulltime, but the estimate remains positive and significant.

Column 7 restricts the sample to graduates who are employed outside the public sector. Private sector wages are presumably more correlated with actual productivity. Thus, if our key hypothesis is true, we would expect to see a larger estimate of the difference-in-difference parameter. And indeed, restricting the sample to private sector workers increases the estimate of the information value of central school exams to 0.057.[13]

Another concern is potential selection into employment. However, as the descriptive statistics have revealed, employment rates in our sample are fairly large (around 90 percent) with no apparent differences between grade levels and types of exit examination. Hence, we refrain from estimating a selection model as our main specification. Instead, we now present results of a specification with employment status as the dependent variable.[14]

Results of the employment regressions are shown in Table 5. Again, different columns refer to specifications with different sets of covariates. As is evident from the key results in the first two rows, both the estimated coefficient of high-school grades alone and the estimated coefficient of the interaction effect are extremely small and insignificant. We interpret these results as reassurance that our difference-in-difference estimates for the effect of high-school grades on earnings are not driven by selection into employment.

So far, we have limited our attention to earnings five years after university graduation and ignored earnings measured in the first wave one year after graduation. One main reason for doing so is that several fields of university studies such as law, medicine, and education involve mandatory post-graduate traineeships that are part of the entire program. These post-graduate traineeships typically last longer than one year and involve a payment of relatively low and identical wages for all participants. Thus, reported earnings one year after graduation do not reflect market wages for these graduates. We can, however, estimate equation (5) with reported earnings one year after graduation for all other graduates.

As shown in Table 6, one year after graduation we obtain an estimate for our key parameter of interest of 0.025 (column 4). While the absolute size of the information value of central school exams is smaller compared to the estimate five years after graduation, the estimated coefficient is positive and significant also one year after graduation. The difference in the size of estimates is, however, driven by the sample restriction, as re-estimating the

[13] In an additional subgroup analysis, we also observe a substantially larger interaction effect for males, while the estimate for females is insignificant and close to zero. Closer inspection of the data, however, reveals that this difference is largely driven by the fact that the majority of females in our sample work in the public sector, while more than two thirds of men work in the private sector.

[14] Despite the binary nature of the dependent variable, we regard our linear model as a good approximation of the conditional expectation function of interest. Estimating a non-linear Probit specification and calculating the cross-derivative following Ai and Norton (2003) produces similar results.

specification five years after graduation with the same sample restriction produces an estimates of 0.023 (column 5).

While the unique institutional setup in Germany together with the information contained in the graduate survey data allow the identification of the information value of central exams, the restriction of the dataset to successful university graduates constitutes a potential drawback. Differences in the selection into the highest track of secondary education, the selection into university studies, and differences in successful graduation between individuals from central and local exam status can hamper identification. In particular, Table 1 already revealed significant differences in track attendance between central and local exam states. Therefore, in the first two columns of Table 7 we test the robustness of our results to additionally controlling for an interaction effect between state-specific cohort shares and grades. The key parameter of interest is positive, although it captures statistical significance only once state fixed effects are included.

Table 1 shows that the state of Bavaria has the lowest cohort graduation share. This may indicate that only the most "able" Bavarian students have a chance to end up in our sample population. To further explore whether sample selectivity could drive our results, in the final two columns of Table 7 we selectively drop observations from other states to match the cohort share of Bavaria. In column 3, we do so by dropping observations at the bottom of the state-specific earnings distributions, while in column 4, we use the state-specific grade distribution. In both cases the estimate for our key parameter of interest remains positive and significant. While we have no formal means of testing to what extent selection constitutes a real concern, at least all additional empirical analyses do not indicate that sample selection drives our results.

7. Characteristics of Universities and Jobs as Channels?

All specifications so far include only control variables that are determined prior to graduation from high school. Additional control variables that characterize university studies or labor-market status are excluded in our main specifications because they might capture part of the effect we are interested in. However, including these variables might shed light on the underlying mechanisms behind the higher information value of central exams. For example, the effect might be driven by selection into different fields of study or occupational sorting. To investigate potential driving factors, we exploit the rich nature of our dataset and estimate specifications including control variables describing university studies and the labor-market status of individuals.

Table 8 reports results based on specifications that add a set of control variables describing the university career of the graduates. In particular, the observed characteristics of the university career include the length of university studies, the final grade obtained at university, 33 indicators for the field of study, and indicators for the federal state of the university. The first four columns include these additional controls individually and the final column includes all of them simultaneously. In all specifications, our key parameter of interest remains positive and significant. Moreover, the size of the effect remains comparable to our baseline estimate. Thus, features of the university careers are unlikely to be a main channel of the information value of central exams.

Table 9 reports results based on specifications that add a set of control variables describing individuals' labor-market status. These include an indicator for working in the public sector, 14 indicators for occupations, 31 indicators for industries, indicators for the federal state of the job, their interaction with grades, and average weekly working hours. Again, in the first six columns these controls are added individually to our baseline specification, while column 7 provides results of a specification that includes all of them simultaneously. Column 8 then additionally adds all university controls included in the final column of Table 8. It is striking to see how robust the main result is. Again, all estimations reveal estimates of the information value of central school exams that are comparable to and do not differ significantly from our baseline estimate.

These results suggest that neither choices of university careers nor occupational choices are a main mechanism by which grades obtained in central high-school exams transform into higher earnings than grades obtained in local exams. It is reassuring that adding the university exam grade to the model does not change the result on the high-school exam grade. Even more, it is particularly reassuring that neither adding fixed effects for the federal state of the university nor adding fixed effects for the federal state of the job affect our results. Quite surprisingly, not even allowing the effect of high-school grades to vary by state of job changes the estimated interaction between central-exam states and high-school grades. Put differently, even for individuals graduating from university in the same state and working on the same state labor market, the fact whether their state of high school had central or local exams has a substantial bearing on the extent to which high-school grades translate into earnings.

8. Evidence from Applications, Interviews, and Job Offers

We close with a brief analysis of differences in performance during the initial job search after graduation from university. We regard the evidence presented in this section as supportive descriptive evidence for a higher information value of central exams.

One year after graduation, survey participants were asked about the number of job applications written, as well as the number of interview invitations and job offers received during their initial job-search period. The evidence presented so far suggests that high-school grades obtained in central as opposed to local exams contain a higher information value. As employers should therefore be more able to judge on effective human capital differences between applicants when grades are obtained from a central rather than local exam, we hypothesize that the type of examination also matters for the performance during the initial job search. In particular, risk-averse employers might *ceteris paribus* be more likely to hire or give interview invitations to applicants with more reliable signals. As a consequence, we speculate that individuals with grades obtained in central exams receive more interview invitations and job offers given the same number of applications.

We test this hypothesis by regressing key characteristics of the initial job-search period on the central exam dummy and a set of other controls. In particular, we focus on the number of applications, interviews, and job offers, as well as interview-per-application and job-offer-per-application ratios as dependent variables. As is evident from the first three columns of Table 10, central exams have no significant association with the number of interviews, applications, or job offers. However, it is possible that individuals with very good grades apply for the job they want the most and get it immediately, whereas bad performers in high school are forced to send out numerous applications and consequently also receive more job offers (but also more rejections). This consideration illustrates that pure numbers of applications, interviews, and offers may not be very informative.

Hence, columns 4 and 5 use the ratios of interviews over applications and job offers over applications, respectively, as dependent variable. As expected, good performers in high school have significantly higher ratios of interviews and job offers per application. More to the point of our analysis, the estimated coefficients on the central-exam dummy are positive and significant. Conditional on grades, high-school graduates with a degree obtained in central as opposed to local exams thus receive more interview invitations and more job offers per application in their initial job-search period. While these associations do not necessarily warrant a causal interpretation and while the specifications do not do full justice to the complexity of the matching process between employers and employees, we think these results provide descriptive evidence that is informative and consistent with an interpretation that central exams have higher information value.

9. Conclusions

We provide first evidence that central school exams have a higher information value than local school exams. Based on a difference-in-difference identification strategy, we find that a one standard-deviation improvement in high-school grades translates into approximately 6 percent higher earnings when grades are obtained in central exams, compared to less than 2 percent when grades are obtained in local exams. The resulting difference of more than 4 percentage points reflects the higher information value of central school exams. Framing the relationship between high-school grades and earnings in a measurement-error model provides a precise interpretation for this difference in estimated coefficients: The higher information value of central school exams reflects the reduction in the "attenuation bias" as grades obtained in central exams are a less noisy signal for effective human capital.

Apart from providing evidence on the quality of sorting on the labor market, this finding closes an important gap in the literature on the impact of institutional structures of the education system on student achievement. While the reduced-form association of central exit examinations with substantially higher learning outcomes of students is much documented, the channels through which the effect operates are less well understood. In theoretical signaling models of educational performance, central exams are hypothesized to reduce the cost and improve the reliability of the information on effective human capital and thus potential productivity to employers, who are then willing to attach higher rewards to better exam outcomes, which in turn increases students' incentives to learn. The argument that central exams increase the extrinsic rewards for learning is, however, an assumption in these models that has not been tested yet. Our findings confirm the association between central exams and higher extrinsic rewards for learning, a necessary condition for the validity of these incentive-based explanations for higher learning outcomes of students when exit exams are central.

References

Ai, Chunrong, and Edward C. Norton. 2003. "Interaction terms in logit and probit models." *Economics Letters* 80, no. 1: 123-129.

Altonji, Joseph G., and Charles R. Pierret. 2001. "Employer learning and statistical discrimination." *Quarterly Journal of Economics* 116, no. 1: 313-350.

Angrist, Joshua D., and Alan B. Krueger. 1999. "Empirical strategies in labor economics." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card. Amsterdam: North Holland: 1277-1366.

Arrow, Kenneth J. 1973. "Higher education as a filter." *Journal of Public Economics* 2, no. 3: 193-216.

Backes-Gellner, Uschi, and Stephan Veen. 2009. "The consequences of central examinations on educational quality standards and labour market outcomes." *Oxford Review of Education* 34, no. 5: 569-588.

Baker, Olesya, and Kevin Lang. 2013. "The effect of high school exit exams on graduation, employment, wages and incarceration." NBER Working Paper 19182. Cambridge, MA: National Bureau of Economic Research.

Becker, William E. 1982. "The educational process and student achievement given uncertainty in measurement." *American Economic Review* 72, no. 1: 229-236.

Becker, William E., and Sherwin Rosen. 1992. "The learning effect of assessment and evaluation in high school." *Economics of Education Review* 11, no. 2: 107-118.

Betts, Julian R. 1998. "The impact of educational standards on the level and distribution of earnings." *American Economic Review* 88, no. 1: 266-275.

Bishop, John H. 1997. "The effect of national standards and curriculum-based examinations on achievement." *American Economic Review* 87, no. 2: 260-264.

Bishop, John H. 2006. "Drinking from the fountain of knowledge: Student incentive to study and learn - Externalities, information problems, and peer pressure." In *Handbook of the Economics of Education, Vol. 2*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 909-944.

Bishop, John H., and Ferran Mane. 2001. "The impacts of minimum competency exam graduation requirements on high school graduation, college attendance and early labor market success." *Labour Economics* 8, no. 2: 203-222.

Bishop, John H., and Ludger Woessmann. 2004. "Institutional effects in a simple model of educational production." *Education Economics* 12, no. 1: 17-38.

Costrell, Robert M. 1994. "A simple model of educational standards." *American Economic Review* 84, no. 4: 956-971.

Daveri, Francesco, and Mika Maliranta. 2007. "Age, seniority and labour costs: Lessons from the Finnish IT revolution." *Economic Policy* 22, no. 49: 117-175.

Dee, Thomas S., and Brian A. Jacob. 2007. "Do high school exit exams influence educational attainment or labor market performance?" In *Standards-based reform and the poverty gap: Lessons for No Child Left Behind*, edited by Adam Gamoran. Washington, D.C.: Brookings University Press.

Donald, Stephen G., and Kevin Lang. 2007. "Inference with difference-in-differences and other panel data." *Review of Economics and Statistics* 89, no. 2: 221-233.

Fitzenberger, Bernd, and Gaby Wunderlich. 2002. "Gender wage differences in West Germany: A cohort analysis." *German Economic Review* 3, no. 4: 379-414.

Grant, Darren. 2007. "Grades as information." *Economics of Education Review* 26, no. 2: 201-214.

Haltiwanger, John C., Julia I. Lane, and James R. Spletzer. 1999. "Productivity differences across employers: The roles of employer size, age, and human capital." *American Economic Review* 89, no. 2: 94-98.

Hanushek, Eric A., and Steven G. Rivkin. 2012. "The distribution of teacher quality and implications for policy." *Annual Review of Economics* 4: 131-157.

Hanushek, Eric A., Guido Schwerdt, Simon Wiederhold, and Ludger Woessmann. 2015. "Returns to skills around the world: Evidence from PIAAC." *European Economic Review* 73: 103-130.

Hanushek, Eric A., and Ludger Woessmann. 2011. "The economics of international differences in educational achievement." In *Handbook of the Economics of Education, Vol. 3*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann. Amsterdam: North Holland: 89-200.

Heckman, James J. 1979. "Sample selection bias as a specification error." *Econometrica* 47, no. 1: 153-161.

Jürges, Hendrik, and Kerstin Schneider. 2010. "Central exit examinations increase performance... but take the fun out of mathematics." *Journal of Population Economics* 23, no. 2: 497-517.

Jürges, Hendrik, Kerstin Schneider, and Felix Büchel. 2005. "The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany." *Journal of the European Economic Association* 3, no. 5: 1134-1155.

Jürges, Hendrik, Kerstin Schneider, Martin Senkbeil, and Claus H. Carstensen. 2012. "Assessment drives learning: The effect of central exit exams on curricular knowledge and mathematical literacy." *Economics of Education Review* 31, no. 1: 56-65.

Kerst, Christian, and Karl-Heinz Minks. 2005. *Fünf Jahre nach Studienabschluss: Berufsverlauf und aktuelle Situation von Hochschulabsolventinnen und Hochschulabsolventen des Prüfungsjahrgangs 1997*. Technical Report 173, HIS-Hochschulplanung. Hannover: Hochschul-Informations-System.

Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the black box: A survey of the matching function." *Journal of Economic Literature* 39, no. 2: 390-431.

Piopiunik, Marc, Guido Schwerdt, and Ludger Woessmann. 2013. "Central school exit exams and labor-market outcomes." *European Journal of Political Economy* 31: 93-108.

Reimer, David, and Jette Schröder. 2006. "Tracing the gender wage gap: Income differences between male and female university graduates in Germany." *Zeitschrift für ArbeitsmarktForschung/Journal for Labour Market Research* 39, no. 2: 235-253.

Spence, A. Michael. 1973. "Job market signalling." *Quarterly Journal of Economics* 87, no. 3: 355-374.

Stiglitz, Joseph E. 1975. "The theory of "screening," education, and the distribution of income." *American Economic Review* 65, no. 3: 283-300.

Woessmann, Ludger. 2003. "Central exit exams and student achievement: International evidence." In *No child left behind? The politics and practice of school accountability*, edited by Paul E. Peterson and Martin R. West. Washington, D.C.: Brookings Institution Press: 292-323.

Woessmann, Ludger. 2005. "The effect heterogeneity of central exams: Evidence from TIMSS, TIMSS-Repeat and PISA." *Education Economics* 13, no. 2: 143-169.

Appendix: A Traditional Difference-in-Difference Setup

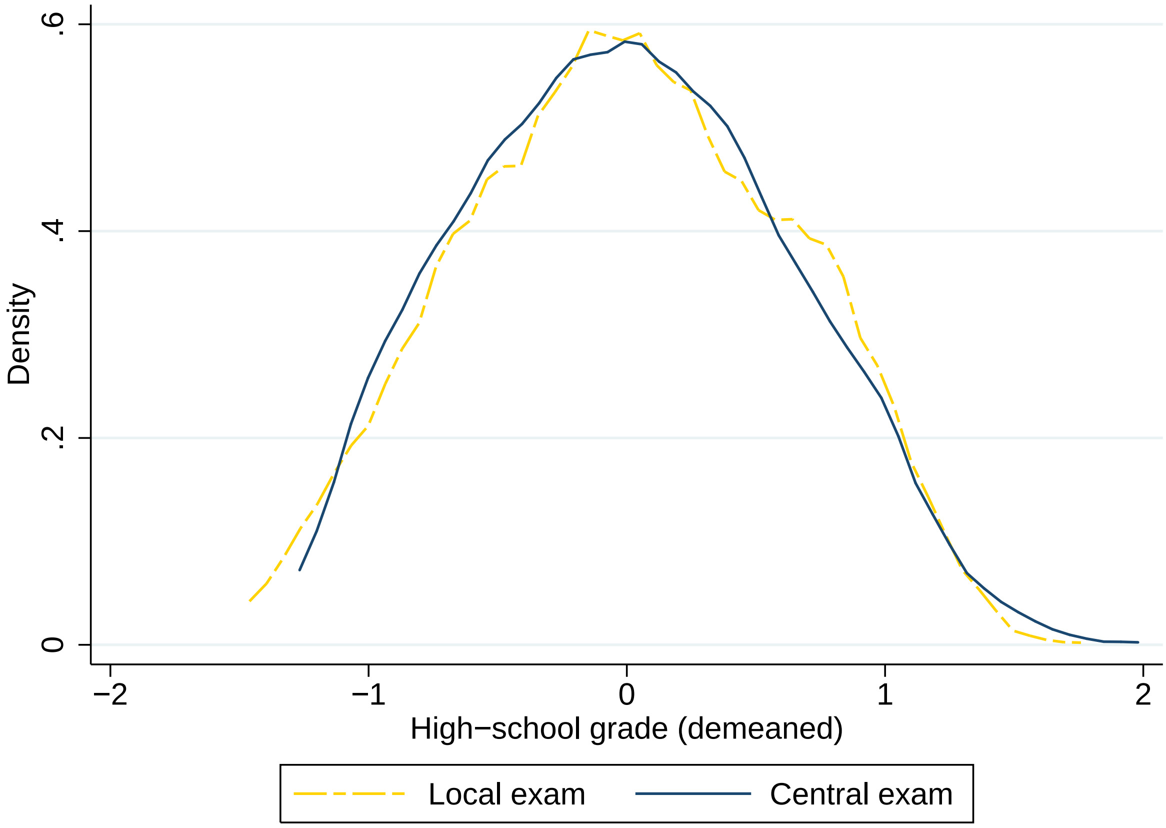
Our baseline specification in equation (5) resembles a difference-in-difference setup. We compare earnings differences between individuals with high and low grades obtained in central exam states to the same difference for individuals who obtained their grades in local exam states. The grade variable used in the empirical analysis is, however, multi-valued. Alternatively, we can estimate a simpler – and possibly more intuitive – specification with a binary measure for school performance based on grades.

Table A-1 reports estimation results based on a simplified version of equation (5) where we substitute the continuous *Grade* variable with a binary indicator, *Good grade*, that indicates whether an individual's grade is better than the state-specific median grade. The results of this estimation confirm the estimates of our baseline specification. Individuals with a good grade in local exams states have on average roughly two percent higher earnings compared to those with a bad grade (although this difference is not statistically significant in this specification). In central exam states, this difference is a significant 6.5 percentage points larger.

While the specification in Table A-1 might be more intuitive, a generic difference-in-difference setup typically includes one difference varying at the level of the observation. In our case, we can construct such a framework by collapsing our sample to include only two observations per federal state. In this case, one observation includes mean characteristics for all individuals with grades below the state-specific median grade and the other observation includes mean characteristics for all individuals with grades above the state-specific median grade.

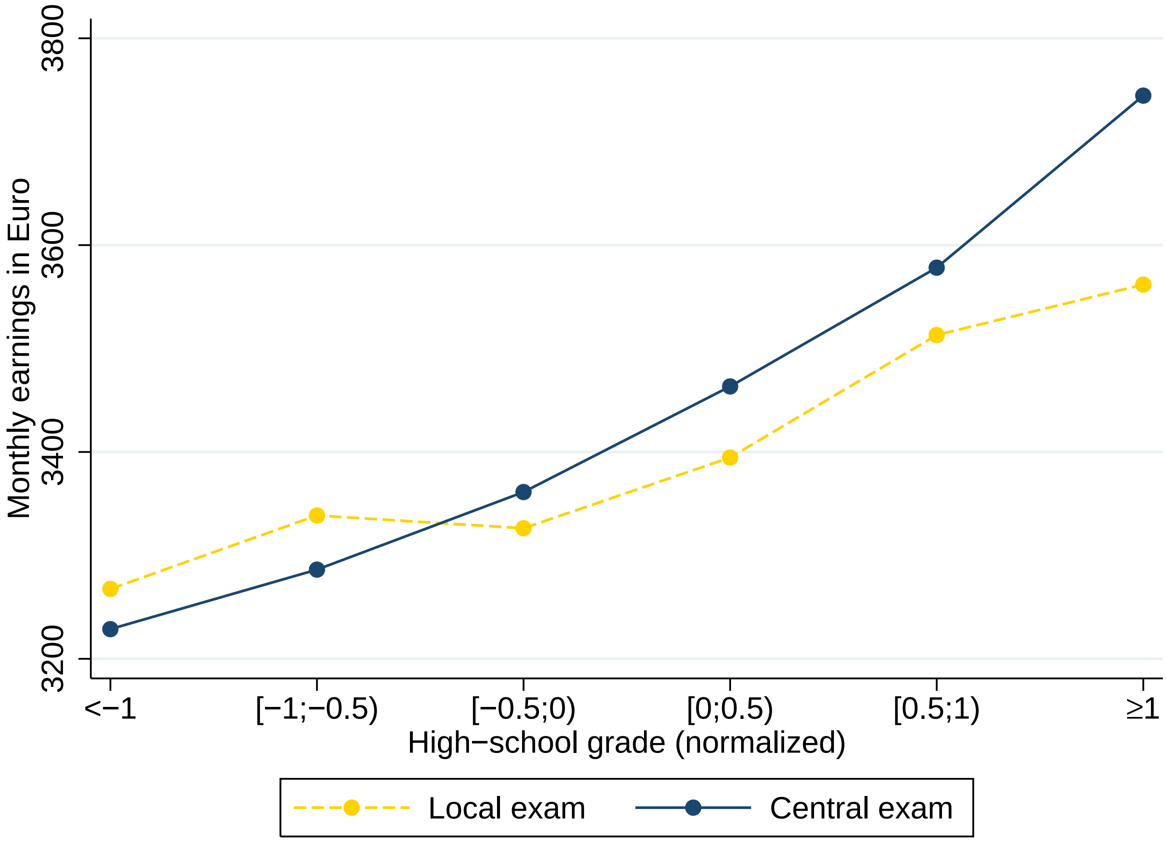
Table A-2 reports results for such a generic difference-in-difference model with 32 observations (two observations per federal state). To account for differences in population size by state, we conduct weighted regressions with the state-specific number of individuals in the original sample over the total number of individuals as weights. The results confirm the findings of Table A-1. In the final two columns, the estimates capture statistical significance at the 10 percent level despite the extremely small sample size.

Figure 1: Distribution of High-School Grades by Examination Type



Note: Kernel density estimates. Grades are demeaned by the respective average grade for each group.

Figure 2: Mean Earnings by High-School Grade and Examination Type



Note: Mean gross monthly earnings in Euro in 2003 by examination type for six intervals based on normalized exam grades.

Table 1: Final Exams at the End of High School in Germany

Local exams			Central exams		
Federal state	Cohort share (1)	Pass rate (2)	Federal state	Cohort share (3)	Pass rate (4)
Berlin	0.31	0.92	Baden-Württemberg	0.20	0.99
Brandenburg	0.28	0.96	Bavaria	0.18	0.99
Bremen	0.31	0.97	Mecklenburg-Vorpommern	0.24	0.97
Hamburg	0.32	0.96	Saarland	0.23	0.98
Hesse	0.27	0.97	Saxony	0.27	0.95
Lower Saxony	0.23	0.96	Saxony-Anhalt	0.25	0.97
North Rhine-Westphalia	0.29	0.96	Thuringia	0.28	0.97
Rhineland-Palatinate	0.21	0.97			
Schleswig-Holstein	0.22	-			
Average	0.29	0.96	Average	0.22	0.98

Note: Share of graduates from the highest school track and pass rate by federal state in Germany. The distinction between central exams and local exams is based on the institutional framework during the 1990s. The cohort shares refer to 1998. The pass rates refer to 2004, the earliest year for which information was available for all federal states.
Source: Federal Statistical Office, Population Statistics and Standing Conference of the Ministers for Education and Cultural Affairs of the *Länder* of the Federal Republic of Germany (KMK).

Table 2: Descriptive Statistics

Variable	Local exams all (1)	Local exams "bad" grade (2)	Local exams "good" grade (3)	Central exams all (4)	Central exams "bad" grade (5)	Central exams "good" grade (6)
Grade	-0.09	-0.88	0.73	0.23	-0.53	1.06
	(1.00)	(0.59)	(0.58)	(0.97)	(0.61)	(0.50)
Earnings	3,390	3,332	3,450	3,476	3,322	3,646
	(1,628)	(1,678)	(1,572)	(1,669)	(1,515)	(1,810)
Age	34.10	34.70	33.47	33.38	33.88	32.83
	(3.00)	(3.11)	(2.74)	(2.77)	(3.14)	(2.17)
Male (share)	0.56	0.57	0.56	0.59	0.60	0.59
West (share)	0.99	0.99	0.99	0.94	0.93	0.94
Fachabitur (share)	0.13	0.14	0.12	0.19	0.23	0.14
Gymnasium (share)	0.79	0.79	0.79	0.74	0.69	0.81
Obs. (employed)	3,120	1,597	1,523	1,581	829	752
Employed (share)	0.89	0.88	0.89	0.89	0.88	0.90
Obs. (all)	3,462	1,795	1,667	1,765	929	836

Note: Means and standard deviations (in parentheses) for selected variables by examination type. Grades are normalized to have mean zero and standard deviation one. "Bad" and "good" grades refer to individuals with below and above median grades within examination types. All variables are measured in 2003. Gross monthly earnings are reported in Euro.

Table 3: Difference-in-Difference Results on the Information Value of Central Exams

	(1)	(2)	(3)	(4)
Central exam*Grade	0.034**	0.041***	0.042***	0.043***
	(0.011)	(0.011)	(0.011)	(0.011)
Grade	0.017**	0.015**	0.017**	0.016**
	(0.007)	(0.007)	(0.007)	(0.007)
Central exam	–0.009			
	(0.036)			
Male	0.407***	0.398***	0.397***	0.397***
	(0.021)	(0.023)	(0.022)	(0.022)
Age	–0.014**	–0.015***	–0.015***	–0.016***
	(0.005)	(0.005)	(0.004)	(0.004)
Fachabitur				–0.038
				(0.024)
Gymnasium				–0.042**
				(0.019)
Constant	8.255***	8.284***	8.258***	8.317***
	(0.180)	(0.163)	(0.173)	(0.167)
High-school state [16]	No	Yes	Yes	Yes
Father's education [6]	No	No	Yes	Yes
Mother's education [6]	No	No	Yes	Yes
Observations	4,701	4,701	4,701	4,701
R^2	0.141	0.151	0.152	0.153

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variable: log gross monthly earnings in 2003. Grades are normalized to have mean zero and standard deviation one. Figures in square brackets indicate the number of included dummy variables. Standard errors clustered at the state level in parentheses.

Table 4: Sub-Sample Analysis

	West	Abitur	Age 18-20	Gymnasium	Usual	Fulltime	Private
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Central exam*Grade	0.037***	0.048***	0.044**	0.038**	0.048**	0.024**	0.057***
	(0.011)	(0.015)	(0.018)	(0.017)	(0.019)	(0.009)	(0.018)
Grade	0.015*	0.018**	0.016*	0.017*	0.014	0.022***	0.031***
	(0.007)	(0.007)	(0.009)	(0.008)	(0.010)	(0.006)	(0.009)
Male	0.405***	0.393***	0.390***	0.405***	0.397***	0.322***	0.476***
	(0.023)	(0.023)	(0.021)	(0.024)	(0.023)	(0.019)	(0.031)
Age	–0.017***	–0.018**	–0.015***	–0.019***	–0.018***	–0.008*	–0.027***
	(0.004)	(0.006)	(0.005)	(0.005)	(0.006)	(0.004)	(0.008)
Fachabitur	–0.029		–0.041	0.069		–0.016	–0.053*
	(0.023)		(0.042)	(0.105)		(0.025)	(0.027)
Gymnasium	–0.032	–0.050**	–0.062*			–0.017	–0.041
	(0.018)	(0.021)	(0.033)			(0.018)	(0.031)
Constant	8.329***	8.397***	8.342***	8.424***	8.411***	8.122***	8.698***
	(0.167)	(0.238)	(0.188)	(0.202)	(0.212)	(0.144)	(0.288)
High-school state [16]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Father's education [6]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother's education [6]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,586	4,005	3,809	3,645	3,356	4,456	3,019
R^2	0.155	0.147	0.153	0.151	0.150	0.141	0.190

* p<0.10, ** p<0.05, ***p<0.01

Note: Dependent variable: log gross monthly earnings in 2003. Column headers indicate the selection criteria for the sub-sample included in the estimation. Grades are normalized to have mean zero and standard deviation one. Figures in square brackets indicate the number of included dummy variables. Standard errors clustered at the state level in parentheses.

Table 5: Employment Regressions

	(1)	(2)	(3)	(4)
Central exam*Grade	0.008	0.007	0.008	0.009
	(0.005)	(0.005)	(0.006)	(0.006)
Grade	0.000	0.001	0.002	0.002
	(0.004)	(0.004)	(0.005)	(0.005)
Central exam	–0.006			
	(0.006)			
Male	0.151***	0.149***	0.148***	0.148***
	(0.009)	(0.009)	(0.009)	(0.010)
Age/10^2	–0.522***	–0.542***	–0.562***	–0.614***
	(0.146)	(0.152)	(0.146)	(0.146)
Fachabitur				–0.020
				(0.023)
Gymnasium				–0.031**
				(0.014)
Constant	0.981***	0.973***	0.965***	1.012***
	(0.053)	(0.053)	(0.053)	(0.046)
High-school state [16]	No	Yes	Yes	Yes
Father's education [6]	No	No	Yes	Yes
Mother's education [6]	No	No	Yes	Yes
Observations	5,227	5,227	5,227	5,227
R^2	0.057	0.061	0.062	0.063

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variable: employment indicator in 2003. Grades are normalized to have mean zero and standard deviation one. Figures in square brackets indicate the number of included dummy variables. Standard errors clustered at the state level in parentheses.

Table 6: Earnings One Year after Graduation

Earnings in	1998				2003
	(1)	(2)	(3)	(4)	(5)
Central exam*Grade	0.016*	0.022**	0.017**	0.025**	0.023**
	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
Grade	–0.016*	–0.017**	–0.011	–0.012	0.011**
	(0.008)	(0.007)	(0.009)	(0.009)	(0.005)
Central exam	0.051				
	(0.034)				
Male	0.314***	0.309***	0.300***	0.285***	0.392***
	(0.024)	(0.023)	(0.023)	(0.024)	(0.030)
Age	–0.007	–0.006	–0.007	–0.009	–0.019**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
Fachabitur				0.114**	–0.041**
				(0.043)	(0.017)
Gymnasium				–0.019	–0.021
				(0.036)	(0.019)
Constant	8.215***	8.201***	8.138***	8.214***	8.456***
	(0.171)	(0.196)	(0.195)	(0.192)	(0.248)
High-school state [16]	No	Yes	Yes	Yes	Yes
Father's education [6]	No	No	Yes	Yes	Yes
Mother's education [6]	No	No	Yes	Yes	Yes
Observations	3,046	3,046	3,046	3,046	3,046
R^2	0.072	0.077	0.092	0.099	0.158

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variable: log gross monthly earnings in 1998 in columns (1) to (4) and log gross monthly earnings in 2003 in column (5). Grades are normalized to have mean zero and standard deviation one. Individuals holding state examination degrees in teaching, law, medicine, and pharmacy are excluded from the sample because these fields involve mandatory post-graduate traineeships. Figures in square brackets indicate the number of included dummy variables. Standard errors clustered at the state level in parentheses.

Table 7: Addressing Sample Selection

	Controlling for cohort share		Truncate to lowest cohort share by earnings	Truncate to lowest cohort share by grades
	(1)	(2)	(3)	(4)
Central exam*Grade	0.034	0.070***	0.019**	0.044**
	(0.023)	(0.016)	(0.007)	(0.016)
Grade	0.019*	0.007	0.017**	0.001
	(0.010)	(0.006)	(0.008)	(0.014)
Central exam	–0.045			
	(0.039)			
Cohort share	–0.535			
	(0.513)			
Cohort share*Grade	0.016	0.376		
	(0.253)	(0.232)		
Male	0.405***	0.398***	0.211***	0.384***
	(0.022)	(0.022)	(0.036)	(0.030)
Age	–0.014**	–0.016***	–0.004	–0.020***
	(0.005)	(0.004)	(0.004)	(0.005)
Fachabitur		–0.037	0.015	–0.041
		(0.024)	(0.020)	(0.025)
Gymnasium		–0.041**	0.032***	–0.034
		(0.019)	(0.010)	(0.022)
Constant	8.398***	8.308***	8.185***	8.440***
	(0.166)	(0.163)	(0.128)	(0.197)
High-school state [16]	No	Yes	Yes	Yes
Father's education [6]	No	Yes	Yes	Yes
Mother's education [6]	No	Yes	Yes	Yes
Observations	4701	4701	3622	3622
R^2	0.142	0.153	0.148	0.145

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variable: log gross monthly earnings in 2003. Grades are normalized to have mean zero and standard deviation one. Figures in square brackets indicate the number of included categories. Cohort share refers to state-specific share of graduates from the highest school track. Columns (3) and (4) truncate each state's sample to match the lowest state's cohort share (0.18), dropping observations at the bottom of the state-specific earnings and grade distribution, respectively. Standard errors clustered at the state level in parentheses.

Table 8: Controlling for Higher-Education Characteristics

	Duration	Exam grade	Field	University state	All
	(1)	(2)	(3)	(4)	(5)
Central exam*Grade	0.044***	0.043***	0.047***	0.039***	0.046***
	(0.012)	(0.011)	(0.010)	(0.011)	(0.010)
Grade	0.014**	0.023***	0.002	0.013**	-0.011
	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)
Male	0.400***	0.398***	0.271***	0.393***	0.264***
	(0.022)	(0.023)	(0.024)	(0.021)	(0.024)
Age	-0.0130***	-0.016***	-0.010**	-0.015***	-0.005
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)
Fachabitur	-0.053**	-0.029	-0.057**	-0.033	-0.048**
	(0.026)	(0.022)	(0.022)	(0.024)	(0.021)
Gymnasium	-0.042**	-0.030	-0.045**	-0.043**	-0.032*
	(0.0191)	(0.020)	(0.016)	(0.019)	(0.017)
Time to degree	-0.008**				-0.016***
	(0.003)				(0.002)
University grade		-0.030***			0.003
		(0.008)			(0.010)
Constant	8.317***	8.320***	7.924***	8.308***	7.923***
	(0.156)	(0.188)	(0.172)	(0.162)	(0.168)
High-school state [16]	Yes	Yes	Yes	Yes	Yes
Father's education [6]	Yes	Yes	Yes	Yes	Yes
Mother's education [6]	Yes	Yes	Yes	Yes	Yes
Field of study [33]	No	No	Yes	No	Yes
University state [18]	No	No	No	Yes	Yes
Observations	4,664	4,531	4,701	4,701	4,499
R^2	0.155	0.156	0.25	0.162	0.262

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variable: log gross monthly earnings in 2003. Column headers indicate the additional controls included in the estimation. Grades are normalized to have mean zero and standard deviation one. Figures in square brackets indicate the number of included dummy variables. Standard errors clustered at the state level in parentheses.

Table 9: Controlling for Occupation Characteristics

	Public	Occupation	Industry	State	State *Grade	Hours	All	All+Uni controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Central exam*Grade	0.046***	0.038***	0.035***	0.043***	0.038**	0.028***	0.030***	0.030***
	(0.012)	(0.010)	(0.008)	(0.010)	(0.014)	(0.007)	(0.009)	(.010)
Grade	0.021***	0.009	0.024***	0.015*	–0.004	0.008	0.017***	-0.002
	(0.007)	(0.007)	(0.004)	(0.007)	(0.049)	(0.006)	(0.005)	(.007)
Male	0.370***	0.331***	0.291***	0.391***	0.391***	0.224***	0.249***	0.213***
	(0.022)	(0.015)	(0.018)	(0.020)	(0.020)	(0.014)	(0.010)	(.011)
Age	–0.015***	–0.013***	–0.010**	–0.014***	–0.014***	–0.008**	–0.006**	-0.003
	(0.004)	(0.003)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)	(.003)
Fachabitur	–0.055**	–0.063**	–0.095***	–0.035	–0.037	–0.010	–0.099***	-0.061***
	(0.020)	(0.022)	(0.018)	(0.023)	(0.024)	(0.026)	(0.017)	(.018)
Gymnasium	–0.040**	–0.051***	–0.044***	–0.039**	–0.041**	–0.043**	–0.049***	-0.037***
	(0.016)	(0.017)	(0.015)	(0.017)	(0.019)	(0.018)	(0.012)	(.012)
Public sector	–0.163***						0.133***	-0.120***
	(0.022)						(0.022)	(.024)
Hours worked						0.023***		
						(0.001)		
Constant	8.358***	8.563***	7.629***	8.195***	8.191***	7.158***	7.912***	8.072***
	(0.165)	(0.121)	(0.296)	(0.172)	(0.163)	(0.133)	(0.225)	(0.184)
High-school state [16]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Father's education [6]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother's education [6]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation [14]	No	Yes	No	No	No	No	Yes	Yes
Industry [31]	No	No	Yes	No	No	No	Yes	Yes
State of job [17]	No	No	No	Yes	Yes	No	Yes	Yes
State of job*Grade	No	No	No	No	Yes	No	No	No
University controls	No	No	No	No	No	No	No	Yes
Observations	4,676	4,666	4,693	4,701	4,701	4,549	4,634	4,403
R^2	0.173	0.274	0.285	0.176	0.178	0.374	0.399	0.422

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variable: log gross monthly earnings in 2003. Column headers indicate the additional controls included in the estimation. Grades are normalized to have mean zero and standard deviation one. Figures in square brackets indicate the number of included dummy variables. University controls refer to the controls for higher-education characteristics of column (5) in Table 8. Standard errors clustered at the state level in parentheses.

Table 10: Evidence on Applications, Interviews, and Job Offers

	Applications	Interviews	Job offers	Interviews/ application	Offers/ application
	(1)	(2)	(3)	(4)	(5)
Central exam	–0.965	–0.116	0.124	0.033**	0.063**
	(1.062)	(0.113)	(0.083)	(0.013)	(0.024)
Grade	–3.857***	–0.119	–0.084**	0.057***	0.027**
	(0.542)	(0.076)	(0.034)	(0.004)	(0.010)
Male	4.659***	0.542***	0.108	0.021	–0.025
	(0.849)	(0.148)	(0.130)	(0.015)	(0.030)
Age	0.143	–0.062***	–0.001	–0.001	–0.003
	(0.177)	(0.016)	(0.019)	(0.002)	(0.004)
Fachabitur	7.023***	0.303	0.022	–0.052**	–0.117***
	(1.534)	(0.265)	(0.256)	(0.018)	(0.028)
Gymnasium	0.304	0.226	–0.085	0.031	0.013
	(0.842)	(0.163)	(0.186)	(0.022)	(0.014)
Constant	9.244	5.145***	2.288**	0.486***	0.561***
	(5.695)	(0.546)	(0.800)	(0.050)	(0.131)
Father's education [6]	Yes	Yes	Yes	Yes	Yes
Mother's education [6]	Yes	Yes	Yes	Yes	Yes
Observations	4,113	3,644	3,598	3,580	3,462
R^2	0.041	0.011	0.005	0.042	0.015

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variables are reported in column headers. Grades are normalized to have mean zero and standard deviation one. Figures in square brackets indicate the number of included dummy variables. Standard errors clustered at the state level in parentheses.

Appendix

Table A-1: Measuring Grades as a Binary Variable

	(1)	(2)	(3)	(4)
Central exam*Good grade	0.065**	0.063**	0.064**	0.066**
	(0.026)	(0.025)	(0.025)	(0.025)
Good grade	0.022	0.022	0.024	0.023
	(0.020)	(0.019)	(0.019)	(0.018)
Central exam	–0.028			
	(0.039)			
Male	0.408***	0.398***	0.398***	0.398***
	(0.022)	(0.024)	(0.023)	(0.023)
Age	–0.015***	–0.016***	–0.016***	–0.017***
	(0.005)	(0.004)	(0.004)	(0.004)
Fachabitur				–0.041*
				(0.023)
Gymnasium				–0.042**
				(0.019)
Constant	8.277***	8.294***	8.270***	8.329***
	(0.165)	(0.147)	(0.157)	(0.151)
High-school state [16]	No	Yes	Yes	Yes
Father's education [6]	No	No	Yes	Yes
Mother's education [6]	No	No	Yes	Yes
Observations	4,701	4,701	4,701	4,701
R^2	0.140	0.149	0.151	0.151

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variable: log gross monthly earnings in 2003. "Good grade" is a dummy variable indicating above-median grades in each state. Figures in square brackets indicate the number of included dummy variables. Standard errors clustered at the state level in parentheses.

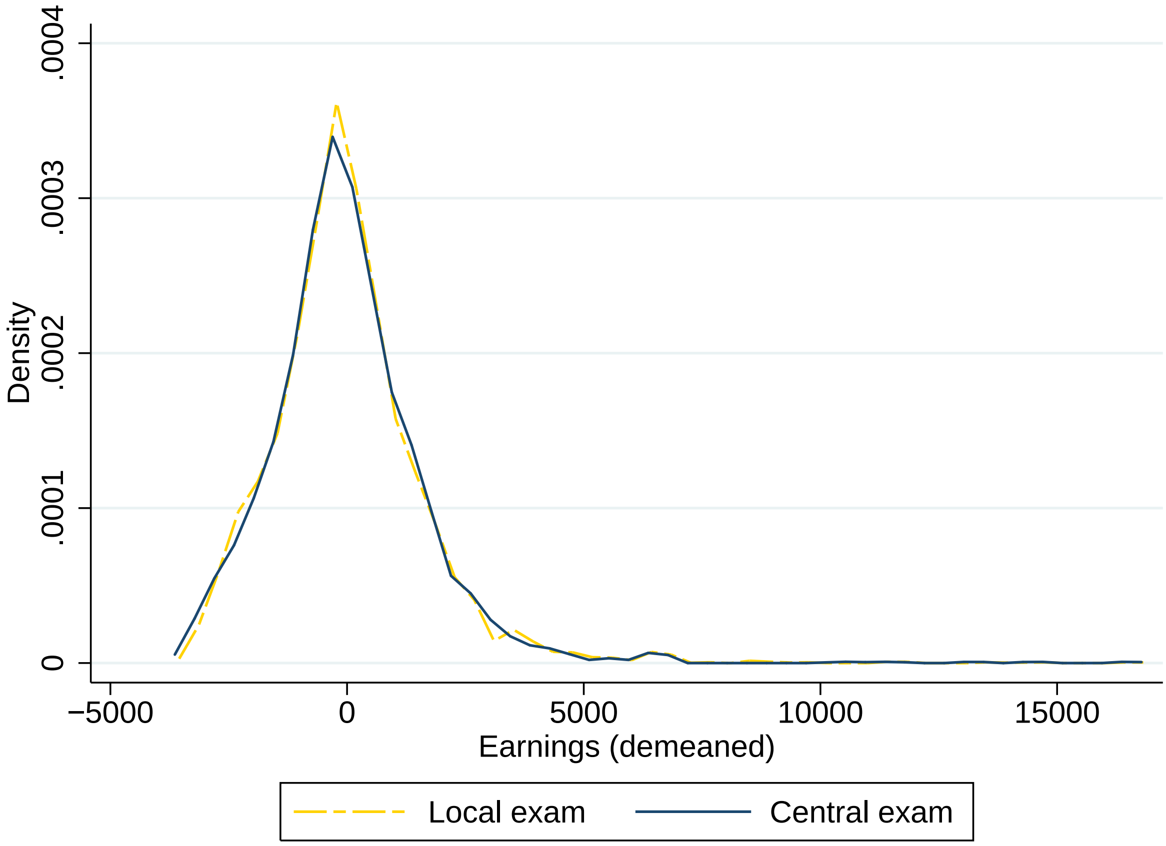
Table A-2: State-Level Estimation: A Generic Difference-in-Difference Model

	(1)	(2)	(3)
Central exam*Good grade	0.063	0.061*	0.068*
	(0.065)	(0.033)	(0.033)
Good grade	0.037	0.038*	–0.027
	(0.037)	(0.019)	(0.056)
Central exam	0.004		
	(0.046)		
Male			0.122
			(0.240)
Age			–0.057
			(0.045)
Constant	7.980***	7.972***	9.893***
	(0.026)	(0.040)	(1.594)
High-school state [16]	No	Yes	Yes
Observations	32	32	32
R^2	0.173	0.892	0.908

* p<0.10, ** p<0.05, *** p<0.01

Note: Dependent variable: log gross monthly earnings in 2003. Coefficients stem from weighted regressions based on the collapsed sample. Weights correspond to the number of observations per federal state in the original sample. "Good grade" is a dummy variable indicating above-median grades in each state. Figures in square brackets indicate the number of included dummy variables. Standard errors clustered at the state level in parentheses.

Figure A-1: Distribution of Earnings by Examination Type



Note: Kernel density estimates. Gross monthly earnings in Euro in 2003. Earnings are demeaned by the respective average earnings for each group.

UNIVERSITY OF KONSTANZ
Department of Economics

Universitätsstraße 10
78464 Konstanz
Germany

Phone: +49 (0) 7531-88-0
Fax: +49 (0) 7531-88-3688

www.wiwi.uni-konstanz.de/econdoc/working-paper-series/

University of Konstanz
Department of Economics