The Signaling Value of Central School Exams

Guido Schwerdt* Ludger Wößmann†

Preliminary Version

Please do not quote without permission.

February 23, 2009

Abstract

While the fact that students perform better in education systems with curriculum-based external exit exams is well documented, little is known about the channels through which this reduced-form pattern arises. One possible channel is that central exams increase the signaling value of high school leaving grades, so that students increase their learning effort to reap later labor-market benefits. This paper tests this channel in a differences-in-differences framework, estimating whether earnings differences between high and low grades are greater in German States without central Abitur exams than ones with them. There is clear evidence that the earnings premium for high school leaving grades is bigger when they are based on central exams. Central exams also increase the number of interviews and job offers per application.

JEL-Code: I20, J31, J24
Keywords: Central exit exams, signaling, earnings, measurement error, differences-in-differences, Germany

---

This research has profited from comments by seminar participants at the ifo Institute in Munich and the Winter Meeting of the North American Econometric Society in San Francisco in 2009. We received insightful comments from Erik Hanushek to whom our gratitude goes.

*CESifo and Ifo Institute for Economic Research, Poschingerstr. 5, 81679 Munich, Germany. Email: schwerdt@ifo.de

†Center for Economic Studies, University of Munich and Ifo Institute for Economic Research, Poschingerstr. 5, 81679 Munich, Germany; Tel.: +49 (89) 9224-1699; woessmann@ifo.de (www.cesifo.de/woessmann).

1 Introduction

Curriculum-based external exit examinations are associated with substantially higher learning outcomes of students (See Bishop, 1997; Wößmann, 2003, 2005; Jürges et al., 2005a,b). While this reduced-form pattern is much documented, the precise channels through which the effect comes about are less well understood. One key mechanism that is often argued to play a major role is that central exams increase the extrinsic rewards for learning by improving the signaling of educational achievement to institutions of higher education and to the labor market (See Bishop, 2006; Bishop and Wößmann, 2004).

Central exams are hypothesized to reduce the cost and improve the reliability of the information (signal) 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. This paper evaluates this channel empirically. We estimate whether the labor market indeed pays higher returns to marks on the high school leaving certificate when the exams that determine these marks are central rather than school specific. We use a unique new data set of German university graduates, part of whom obtained their high school grades in states with central exams and part of whom in states without central exams. By surveying their high school grades and following them into the labor market, the data set for the first time allows to estimate whether the signaling value of central exams is higher than the signaling value of local grades.

This paper builds on the theoretical literature on signaling educational performance.[1] In particular, Bishop and Wößmann (2004) argue that central examinations change the 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 schooling quality when they make admissions and hiring decisions. In consequence, their decisions become less sensitive to other factors such as family connections, the chemistry of a twenty-minute job interview, performance relative to the class mean, or aptitude tests which lean more to measuring innate ability than to measuring overall educational performance. As students' rewards for learning grow, students respond by increasing their learning effort. The idea that central exams increase the extrinsic rewards of learning is, however, an assumption in their model that has not been tested yet.

This paper tests this necessary condition for the signaling explanation to work. The

[1] See Spence (1973), Stiglitz (1975) and Arrow (1973) for seminal papers in this literature.

identification strategy of this paper is not based on variation in earnings levels between states with different examination types, which are differenced out, but based on comparing earnings of high and low markers depending on whether they obtained the marks in central or local exam states. The identification strategy thus resembles a difference-in-difference setup. Based on cross-sectional data we estimate an earnings function with state fixed effects and focus at the interaction of a dummy for central exam states with high school grades. The hypothesis is that this coefficient should be negative.[2]

The remainder of the paper is structured as follows: The next section provides a simple framework for the signaling value of central exams on the labor market. Section 3 describes the examination systems at the end of high school in the German States, introduces the data set, describes the sample selection and provides descriptive statistics. Section 4 presents the econometric strategy. Sections 5 and 6 report the basic results and several robustness tests. Section 7 presents supporting evidence on a higher signaling value of central exams based on applications, interviews, and job offers during the initial job search. Section 8 concludes.

2 Signaling Educational Performance

2.1 Theoretical background

This study builds on the theoretical literature on signaling educational performance. The idea of job market signaling was introduced in the seminal work of Spence (1973): High ability workers differentiate themselves from observationally identical workers of lower ability by acquiring an educational signal which 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 in Stiglitz (1975). In an economy with imperfect information, heterogeneous jobs demanding less or more 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. In his model college education serves as a screening device, in that it sorts out individuals of differing abilities

[2]We expect the coefficient to be negative, because grades in German high schools range from 4.0 to 1.0 with 1.0 being the best possible grade.

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). In this paper the screening model of Stiglitz (1975) is extended by explicitly linking post-school income to the assessment of the college of a student's academic achievement and the prospective employer's assessment 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. In another theoretical paper, 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 was also chosen in studies that more directly focus on the relationship between educational standards and earnings. Costrell (1994, 1997) and Betts (1998) analyze the optimal setting of educational standards when students from school 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 higher social welfare than decentralized standard setting.

The theoretical framework closest to this paper is Bishop and Wößmann (2004). In their model central examinations change the 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 schooling quality when they make admissions and hiring decisions. In consequence, their decisions become less sensitive to other factors such as family connections, the chemistry of a twenty-minute job interview, performance relative to the class mean, or aptitude tests which 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, transition to the institution of central examinations central exams should have a positive effect on the rewards for learning, especially on the extrinsic part. As students' rewards for learning grow, anything which increases the quality of schooling 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 schooling quality. However, the idea

that central exams have a higher signaling value is an assumption in their model. This assumption is tested in this paper.

2.2 Central School Exams in a Measurement Error Framework

The signaling value of central school exams can be expressed in a measurement error framework. For illustrative purposes 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 (signal) 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 form central exit exams carry less noise than others. 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:

$$g_i^c = a_i + n_i \tag{1}$$

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:

$$g_i^{nc} = a_i + n_i + e_i \tag{2}$$

This setup allows us to define the signaling 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:

$$w_i = a_i * \beta + \epsilon_i \tag{3}$$

Assuming that the measurement error, e_i, and the equation error ϵ_i are uncorrelated, the regression of wages on high school grades for graduates with non-central exit examinations can be expressed as:

$$w_i = g_i^{nc} \lambda \beta + \widetilde{\epsilon_i} \tag{4}$$

, where $\lambda = C(g^{nc}, a)/V(g^{nc})$. If high school grades for graduates with non-central exit exams proxy for productive skills with classical measurement error, then $C(g^{nc}, a) = V(a)$

and $V(g^{nc}) = V(a) + V(e)$, so the regression coefficient is necessarily attenuated, with the proportional "attenuation bias" equal to $(1 - \lambda) < 1$.

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

Altonji and Pierret (2001) show that with the passage of time employers learn about the true productive skills of a worker. In our setup this can happen in two ways: First, employers learn by observing on-the-job productivity. Hence, with tenure increasing the importance of the signal diminishes. Second, workers have acquired other signals after graduating from high school. In our empirical setup all workers have additionally obtained a university degree. This clearly is another signal of productive skills. Hence, wage setting in the first job will also be based on other information on effective human capital.

In the limit case of complete employer learning wages will be set independently of any signal of educational achievement. However, as long as employer learning is not instantaneous workers suffer from short run cost of obtaining a less reliable signal. Imagine a situation where employers have no other information on productive skills of a worker other than the high school grade. A risk-neutral, profit maximizing employer will price labor inputs according to the observed relationship between high school grades and productivity.[3] Hence, wage offers in this situation depend on high school grades, the information about the exam type of the high school exam and estimates for β. As shown above, the estimate for β is attenuated towards zero in the case of local exams. Hence, even in the presence of employer learning students in central exams regimes have a higher incentive to obtain a better grade due to higher extrinsic rewards in the short run.

[3]It seems realistic to assume that employers understand the signaling differences between central and local grades. In a multiple generations model of the labor market, employers could have learned about the relation between grades obtained in either central or local exams and productivity by observing this relation in previous generations at the time when employer learning was complete.

3 The German Examination Systems and the HIS Data

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

Young adults in Germany take at the end of their secondary education, usually after 12 or 13 years of schooling, final exams in different subjects. 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*.[4] 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.[5] 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 the final exams. 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 regions in Germany.

Most importantly for this study, 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 examination 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 in the early 1990's. Table 1 provides an overview of the type of exit examination by federal state. The table shows that by the mid 1990s 7 states had introduced external exit examinations, while 9 states hadn't. This variation in institutional environments constitutes the main ingredient of our identification strategy.

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

[5] 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*.

3.2 The HIS Data

Since 1989, selected cohorts of university graduates (1989, 1993, 1997) have been included in large representative surveys conducted by the HIS Ltd. The results were published in several field studies (e.g. Kerst and Minks (2005) and other reports published by HIS). In this study we use data for the 1997 cohort only. University graduates of 1997 were surveyed in a first wave about 1 year after graduation (1998) and again in a second wave about 5 years after graduation (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 9583 graduates were initially sampled in 1998 and 6220 individuals were sampled again in 2003. Our data set in hand contains information on 6216 graduates, who were sampled in both waves.[7] This corresponds roughly to 3.6 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 had to be dropped due to these sample restrictions.

The data set 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)[8] and the federal state in which they obtained their *Abitur*. This enables us to identify whether the final exams were external or local exit examinations. Moreover, we can link this information on high school grades and examination type to labor-market outcomes five years after graduation from university.

3.3 Descriptive Statistics

Compositional differences between groups also cause variation in labor market outcomes. Hence, 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 we distinguish in table 2 only between individuals with good and

[6] The data was provided by the Gesis-ZA Zentralarchiv für Empirische Wirtschaftsforschung. The data became available as a scientific use file (ZA 4272) in 2007.

[7] 4 observations had to be dropped due to anonymity concerns.

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

bad grades. We consider a grade to be a good grade if the grade is below the median grade given a specific examination type and vice versa. The entries in the table represent mean values and standard deviations (in parenthesis).

Comparing mean values for age, the share of males and the share of individuals graduating in West-German states reveals only small differences between groups. Somehow 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% in the non-central exam group compared to 19% 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 p.p. among central exam graduates. Furthermore, there exists a notable difference of 12 p.p. 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.

Analyzing wage differentials between groups requires taking into account potential differences in labor supply as selection into employment might partially drive results.[9] Table 2 therefore reports also group specific employment rates and hours worked.[10] 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 rigourously in the robustness section.

Finally, table 2 provides information on differences in average grades and earnings between groups. The average grade among graduates from central exam states, 2.1, is slightly better than the average grade in the non-central exam group, 2.3. More importantly, mean differences between the groups of individuals with good and bad grades are identical with exactly one entire grade level. 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 are also associated with on average higher earnings in 2003 as opposed to local exam graduates. The difference in monthly earnings is roughly 100 EUROs. This difference might not be surprising given the mean difference in grades. What is, however, more surprising is the within examination type

[9] See the seminal work of Heckman (1979).

[10] Participants in the survey were asked about the amount of hours worked usually worked in their main job including paid and unpaid overtime.

difference in earnings between graduates with above and below median grades. 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.

Figure 2 makes this point even more obvious by plotting average monthly earnings against grade categories separately for central exam and non-central exam states. Figure 2 provides eye-ball evidence for a difference in the correlation between high school grades and earnings by type of examination. While only a slight 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 external exit examination regime is high.

4 The Econometric Model

We estimate the differential effect of high school grades on earnings by type of examination within a difference-in-difference framework. The empirical model is defined as follows:

$$y_i = \alpha + \beta Grade_i + \gamma Grade_i \times CenExam_i + \delta X_i + \epsilon_i \tag{5}$$

where y_i represents log monthly earnings, $Grade_i$ is the high school grade, $CenExam_i$ is an indicator variable for grades obtained in central exams, X_i is a vector of personal characteristics including a set of dummy variables identifying the federal state of the exit examination and, ϵ_i, is an error term uncorrelated with all right-hand side variables. In equation (5) the isolated effect of high school grades on earnings is captured by the parameter β. In our specification β reflects the effect of high school grades on earnings in states with local exit examinations. The isolated effect of curriculum-based external exit examinations is subsumed by the inclusion of a set of dummy variables indicating the federal state of the exit examination.

The key parameter of interest, γ, is identified by the estimate for the interaction effect between $Grade_i$ and $CenExam_i$. The parameter estimate γ captures any additional effect of the high school grade on earnings when grades are obtained in an external examination that goes beyond the isolated effect of high school grades on earnings. In light of the measurement error model presented in section 2 the coefficient estimate for the parameter γ can be interpreted as the reduction of the "attenuation bias" due to the less noisy signal of high school grades obtained in central exams.

Our identification strategy rests on the comparison between graduates from central

exam states and graduates from non- central exam states. It rests on the assumption that, conditional on other control variables, 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 central exam states graduates 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 Results

This section reports results from regressing log monthly earnings on high school grades, high school grades interacted with a central exam dummy ($CenExam$) and different sets of other explanatory variables. Results from estimating different specifications of equation (5) are presented in table 3. All specifications (but the first) include state fixed effects, a male dummy, age in 2003 and a constant term. All estimations allow for clustered standard errors by federal state.[11]

We discuss first the estimated coefficients of the other control variables. Table 3 shows that the estimated coefficients for the male dummy are positive and significant in all specifications. This finding is in line with other evidence on the existence of a gender wage gap in Germany (See Fitzenberger and Wunderlich, 2002; Reimer and Schroder, 2006). The negative and significant negative effect of age seems striking at first glance. Most empirical studies investigating the age-productivity profile find a concave pattern suggesting a positive age effect for early and mid career workers (See Daveri and Maliranta, 2007; Haltiwanger and Spletzer, 1999). However, note that by focusing solely on 1997 university graduates we implicitly introduce a restriction for potential work experience after graduation. Moreover, those graduating at an older age might have needed more time to obtain their degree, which suggests that selection issues might be the explanation for the negative coefficient for our age variable.

All specifications but the first include state fixed effects. This set of dummy variables for the state of the high school captures all grade-invariant variation between states. In the

[11] 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 as recently shown by Donald and Lang (2007). However, all our results hold also without clustering.

presence of mobility costs, comparison of labor markets outcomes of high school graduates from states with and without central exams might be biased due to differences in local labor market conditions. The most parsimonious specification presented in the first column includes a dummy indicating a state with central examination system instead of state fixed effects. The positive coefficient on the $CenExam$ dummy reveals that central exam states are associated with on average higher earnings. This potentially reflects better local labor market conditions in states with central exams. Existing evidence, however, suggests that 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, 2008). An alternative explanation for the positive coefficient would be that central exams have true productivity implications through their positive impact on student learning. While this indeed appears to be a plausible consequence of increased student learning triggered by central exams, it is not the primary concern of this paper.

In fact, in all other specifications we difference out any systematic differences between states by including state fixed effects. Moreover, the specifications in the last 4 columns control also for other confounding factors. Column 3 introduces controls for family background, column 4 for the type of high school degree, column 5 for the degree issuing institution and column 6 for hours worked. Introducing these controls does not affect the other coefficient estimates. Only the introduction of hours worked has a somehow stronger impact. Controlling for differences in hours worked lowers the estimated effect of both grades and grades interacted with the central exam dummy. However, hours worked might be also directly affected by high school grades and the type of exit exam. In this case, controlling for hours worked would lead to a downward bias. We, nevertheless, include hours worked as a control to show that our main effect is not driven by variation in labor input.

The main effect of interest is presented in the first two rows of table 3. The point estimate for the effect of grades on earnings is negative and significant and varies between -.03 and -.012 according to specification. As grades are coded between 1.0 and 4.0 in Germany, with 1 being the best and 4 being the worst grade, the coefficient can be interpreted as follows: a better grade by one entire grade level translates into .03 to .012 percent higher earnings. It represents the baseline effect of grades on earnings in states with local grades. In light of the measurement framework presented in section 2 this can be interpreted as a biased estimate for the effect of high school grades on earnings.

Row 2 of table 3 shows that this effect is significantly different in states with central

exams. Here the association between high school grades and future earnings is much stronger. The estimates for the interaction effect range from -.056 to -.042 suggesting that a by one grade level better grade translates into .075 to .055 percent higher earnings in central exam states. The estimated interaction effects can be 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 external exit exams. We call this parameter estimate the signaling value of central school exams.

6 Sensitivity Analysis

This section provides evidence on the robustness of our results. One 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 repeating the above analysis for different subgroups of the population.

Table 8 provides regression results based on equation (5) for five different subgroups of our sample. All results are based on the most elaborate specification of equation (5), which corresponds to column 6 of table 3. The first column of table 8 shows estimated coefficients based on individuals from the former West German states only. The results remain qualitatively unchanged with a somehow lower point estimate for the interaction effect of grades and the $CenExam$ dummy.

Columns 2 to 4 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 18-20 and column 4 to individuals who obtained their degree at a *Gymnasium*. Again, our key result of interest remains robust. The magnitude of the estimates for the signaling value of central school exams becomes even slightly larger.

Finally, column 5 presents results based an estimation sample for individuals working fulltime in 2003. Differences in the selection into fulltime employment by grade level and federal state may significantly affect the estimates of our baseline specification. However, column 5 reveals that our key results hold even when focusing on those individuals working fulltime.

Selection into employment might be an even more important concern. However, as the

descriptive statistics have revealed, employment rates in our sample are fairly large (around 0.9) and no significant differences between grade levels and type of exit examination are apparent. Hence, we refrained from estimating a selection model as our main specification. Instead we now present results from a Probit estimation of employment status on high school grades, a central exam dummy, high school grades interacted with the central dummy and other covariates. This estimation resembles the first stage of a Heckman-style selection model.

The results of this exercise can be seen in table 5. Again, the results presented in different columns refer to different specifications including different sets of covariates. The key results of interested are reported in row 1 and 2. While the estimated marginal effects for high school grades alone are all negative and highly significant, the estimates for the interaction effect are extremely small in magnitude and insignificant. Only the marginal effect in column 3 is significant at the 10% level. 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.

7 Evidence from Applications, Interviews, and Job Offers

This section discusses 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 signaling value of central exams.

One year after graduation participants in the survey were asked about the number of applications written as well as 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 signaling value. As employers should therefore be more able to judge on relative effective human capital differences between applicants when grades are obtained in a central exam, one can 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 the more reliable signal. As a consequence, one can speculate that individuals with grades obtained in central exams receive more interview invitations and job offers given the same amount of applications.

We test this hypothesis by regressing key characteristics of the initial job search period

on a $CenExam$ 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. The results of this exercise can be seen in table 6. Column 1 to 3 report results with number of applications, interviews and job offers as dependent variables. While central exams have a positive, significant impact only on the number of job offers, high school grades have a positive, significant effect on the number of applications and the number of job offers. Note that a positive effect of grades indicates that individuals with worse high school grades are associated with more applications and also more job offers. One can speculate that individuals with very good grades apply for the job they want the most and get it immediately, while 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 are not the most interesting statistics to look at. Hence, we present in columns 4 and 5 evidence on the ratios of interviews over applications and job offers over applications, respectively. Estimated coefficients indicate that good performers in high school are associated with a significantly higher interviews- and job offers per application ratios, which is what we would expect. More importantly for this analysis are the estimated coefficients on the central exam dummy. They are both positive and significant. Hence, conditional on grades high school graduates with central exams as opposed to those with a degree obtained in local exams are also associated with better interviews- and job offers per application in their initial job search period.

We regard these results as supporting descriptive evidence on the higher signaling value of central exams. Note that we are well aware that our specifications presented in table 6 are very simplistic and don't do justice to the complexity of the matching process between employers and employees. Moreover, one might be worried about variation between federal states (differences in local labor markets for example) that potentially bias our estimates. Consequently, these results should not be interpreted causally. We nevertheless regard these results as informative as the key estimates of interest have the expected signs. The estimated positive relation between obtaining high school grades in a central exam and a better performance during the initial job search is consistent with our presumption that central exams produce more reliable signals of individual human capital and therefore also positively affect job search outcomes.

8 Conclusion

We provide first evidence on a higher signaling value of central school exams. Based on a difference-in-difference identification strategy we show that a grade improvement by one entire grade level translates into approximately 8 percent higher earnings when grades are obtained in central exam compared to a 3 percent increase in earnings when grades are obtained in local exams. The resulting difference of 5 percentage points reflects the higher signaling value of central school exams. Framing the relation between high school grades and earnings in a measurement error model, provides a precise interpretation for this difference in estimated coefficients: the higher signaling value of central school exams reflects the reduction in the "attenuation bias" as grades obtained in central exam are a less noisy signal for effective human capital.

This finding closes an important gap in the literature on the impact of institutions of the education system and student achievement. While the reduced-form association between central exit examinations with substantially higher learning outcomes of students is much documented, the precise channels through which the effect comes about are less well understood. In the theoretical model of Bishop and Wößmann (2004) central exams are hypothesized to reduce the cost and improve the reliability of the information (signal) 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 of learning is, however, an assumption in their model that has not been tested yet.

For the first time this paper tests this channel empirically. The findings confirm the association between central exams and higher extrinsic rewards of learning. This can be seen as a necessary condition for the validity of the signaling explanation for substantially higher learning outcomes of students when exit exams are central.

References

Ai, C. and Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, 80(1):123–129.

Altonji, J. G. and Pierret, C. (2001). Employer learning and statistical discrimination. *Quarterly Journal of Economics*, 116:313–350.

Arrow, K. J. (1973). Higher education as a filter. *Journal of Public Economics*, 2(3):193–216.

Backes-Gellner, U. and Veen, S. (2008). The consequences of central exams on educational quality standards and labour market outcomes. *Oxford Review of Education*, 43(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 Rosen, S. (1992). The learning effect of assessment and evaluation in high school. *Economics of Education Review*, 11(2):107–118.

Betts, J. R. (1998). The impact of educational standards on the level and distribution of earnings. *American Economic Review*, 88(1):266–75.

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

Bishop, J. H. (1997). The effect of national standards and curriculum-based exams on achievement. *American Economic Review*, 87(2):260–64.

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

Costrell, R. M. (1994). A simple model of educational standards. *American Economic Review*, 84(4):956–71.

Costrell, R. M. (1997). Can centralized educational standards raise welfare? *Journal of Public Economics*, 65(3):271–293.

Daveri, F. and Maliranta, M. (2007). Age, seniority and labour costs: lessons from the finnish it revolution. *Economic Policy*, 22:117–175.

Donald, S. G. and Lang, K. (2007). Inference with difference-in-differences and other panel data. *The Review of Economics and Statistics*, 89(2):221–233.

Fitzenberger, B. and Wunderlich, G. (2002). Gender wage differences in west germany: A cohort analysis. *German Economic Review*, 3(4):379 – 414.

Haltiwanger, John, J. L. and Spletzer, J. (1999). Productivity differences across employers. the roles of employer size, age and human capital. *American Economic Review*, 89(2):94–98. Papers and Proceedings.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–61.

Jürges, H., Richter, W. F., and Schneider, K. (2005a). Teacher quality and incentives: Theoretical and empirical effects of standards on teacher quality. *FinanzArchiv: Public Finance Analysis*, 61(3):298–326.

Jürges, H., Schneider, K., and Büchel, F. (2005b). 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.

Kerst, C. and Minks, K.-H. (2005). Fünf jahre nach studienabschluss. berufsverlauf und aktuelle situation von hochschulabsolventinnen und -absolventen des prüfungsjahrgangs 1997. Technical Report 173, HIS-Hochschulplanung.

Reimer, D. and Schroder, J. (2006). Tracing the gender wage gap: Income differences between male and female university graduates in germany. *Zeitschrift für Arbeitsmarkt Forschung/Journal for Labour Market Research*, 39(2):235 – 253.

Spence, A. M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–74.

Stiglitz, J. E. (1975). The theory of screening, education, and the distribution of income. *American Economic Review*, 65(3):283–300.

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

Wößmann, L. (2005). The effect heterogeneity of central examinations: evidence from timss, timss-repeat and pisa. *Education Economics*, 13(2):143–169.

Tables & Figures

Table 1: Introduction of central exams in Germany

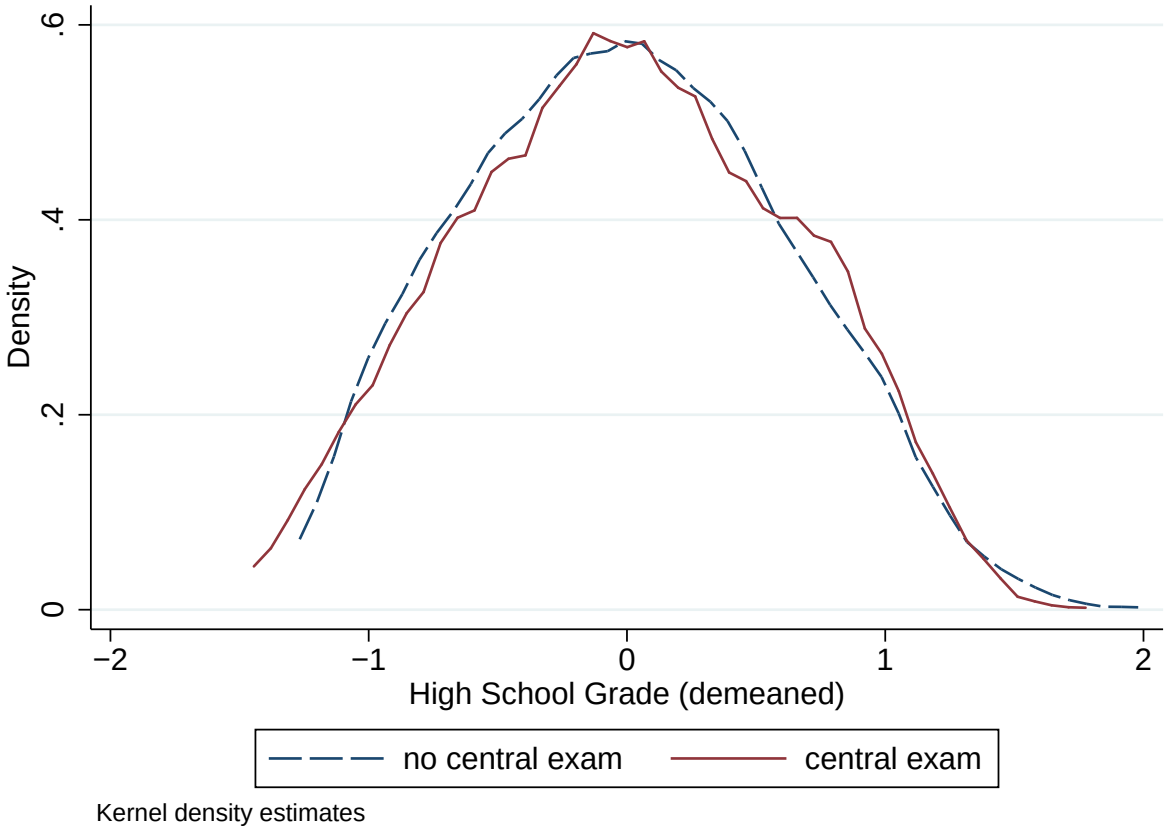
Federal State	**central exam introduced in year**	
Rhineland-Palatinate	not introduced	
Schleswig-Holstein	2008	(not for all subjects)
Bremen	2007	
North Rhine-Westphalia	2007	
Hesse	2007	
Berlin	2006	
Lower Saxony	2006	(not for all subjects)
Hamburg	2005	(not for all subjects)
Brandenburg	2005	(not for all subjects)
Saxony	1993	
Saxony-Anhalt	1993	
Mecklenburg-Vorpommern	1991	
Thuringia	1990	
Baden-Württemberg	1952	
Bavaria	1946	
Saarland	1945	

Table 2: Descriptive statistics

Variable		local exam			central exam	
	all	bad grade	good grade	all	bad grade	good grade
Age	34	35	33	33	34	33
	(3.1)	(3.1)	(2.9)	(2.8)	(3.1)	(2.3)
Male	.55	.55	.54	.59	.6	.57
(share)	(.5)	(.5)	(.5)	(.49)	(.49)	(.49)
West	.96	.97	.96	.94	.93	.95
(share)	(.18)	(.18)	(.19)	(.24)	(.25)	(.23)
Fachabitur	.13	.14	.12	.19	.24	.13
(share)	(.34)	(.35)	(.33)	(.39)	(.43)	(.34)
Gymnasium	.78	.77	.79	.75	.69	.81
(share)	(.41)	(.42)	(.41)	(.43)	(.46)	(.39)
Employed	.9	.89	.92	.9	.89	.9
(share)	(.3)	(.31)	(.28)	(.31)	(.31)	(.3)
Hours	44	43	44	44	43	45
	(12)	(11)	(12)	(11)	(12)	(11)
Grade	2.3	2.8	1.8	2.1	2.6	1.6
	(.63)	(.39)	(.37)	(.61)	(.39)	(.32)
Earnings	3,370	3,316	3,426	3,478	3,323	3,651
	(1,615)	(1,665)	(1,561)	(1,671)	(1,516)	(1,813)

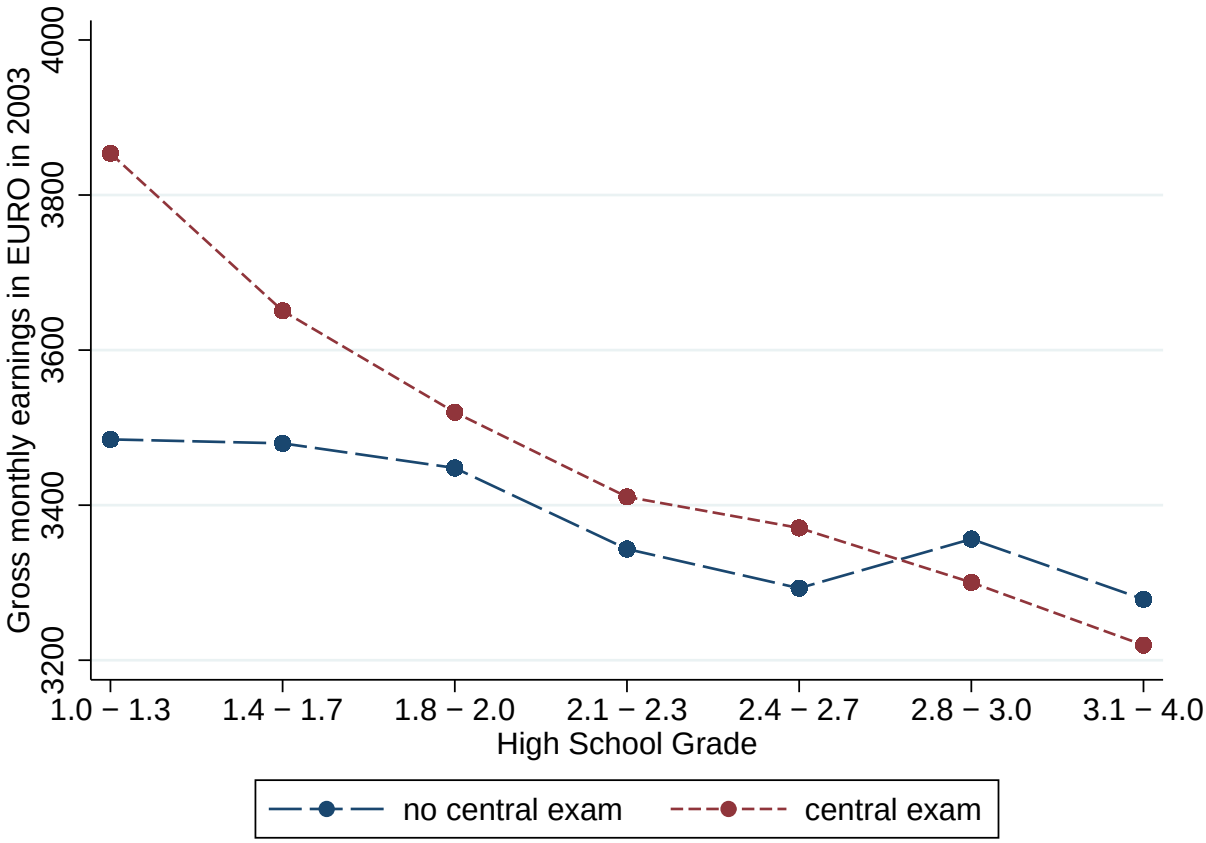
Note: The table contains mean values and standard deviations (in parenthesis) for selected variables by examination type. Every second row reports mean values separately for individuals with above and below median grades. All variables are measured in 2003. Monthly gross earnings are reported in EUROs.

Figure 1: Distribution of high school grades by type of examination



Note: Grades are demeaned by the respective average grade for both groups.

Figure 2: Mean earnings by grade and examination type



Note: Grades are pooled in 6 categories according to the German grade classification scheme with a different grade label every .3 grade point. Grades below 3 are pooled into one category to ensure comparable numbers of observations within categories.

Table 3: Estimation Results

	1	2	3	4	5	6
CenExam*Grade	–.0593**	–.0551**	–.0572**	–.0575**	–.0583**	–.0435**
	(.019)	(.014)	(.014)	(.013)	(.013)	(.010)
Grade	–.0216*	–.0283**	–.0298**	–.0298**	–.0288**	–.0126
	(.012)	(.010)	(.011)	(.011)	(.011)	(.009)
CenExam	.134**					
	(.056)					
Male-dummy	.406**	.396**	.396**	.395**	.395**	.221**
	(.021)	(.023)	(.022)	(.022)	(.022)	(.014)
Age	–.0131**	–.0153**	–.0153**	–.0154**	–.0160**	–.00818**
	(.005)	(.005)	(.004)	(.004)	(.004)	(.003)
Fachabitur				.00486	–.0274	.00432
				(.015)	(.025)	(.029)
Gymnasium					–.0389*	–.0317
					(.019)	(.020)
Hours						.0230**
						(.001)
Constant	8.265**	8.350**	8.326**	8.328**	8.383**	7.189**
	(.200)	(.168)	(.173)	(.174)	(.166)	(.119)
Father's education	No	No	Yes	Yes	Yes	Yes
Mother's education	No	No	Yes	Yes	Yes	Yes
State of High School	No	Yes	Yes	Yes	Yes	Yes
Observations	4829	4829	4801	4801	4801	4640
R^2	.139	.152	.153	.153	.153	.372

* p<0.10, ** p<0.05

Note: Dependent variable is log monthly earnings in 2003. Clustered standard errors in parenthesis.

Table 4: Subgroup Analysis

	WEST	ABI	Age 18-20	Gym.	Usual scen.	Fulltime
CenExam*Grade	–.038**	–.048**	–.053**	–.050**	–.068**	–.032**
	(.010)	(.011)	(.015)	(.014)	(.017)	(.011)
Grade	–.012	–.0089	–.0043	–.0053	.0024	–.021**
	(.009)	(.009)	(.011)	(.009)	(.011)	(.009)
Male-dummy	.23**	.22**	.22**	.23**	.22**	.22**
	(.015)	(.018)	(.016)	(.017)	(.017)	(.014)
Age	–.0087**	–.0089*	–.0077	–.0095**	–.0095*	–.0054*
	(.003)	(.004)	(.005)	(.004)	(.005)	(.003)
Fachabitur	–.00084		.0041	.047		.0062
	(.026)		(.045)	(.089)		(.026)
Gymnasium	–.032*	–.036*	–.035			–.014
	(.016)	(.020)	(.027)			(.018)
Hours	.023**	.023**	.023**	.024**	.024**	.015**
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Constant	7.21**	7.21**	7.17**	7.17**	7.20**	7.48**
	(.122)	(.160)	(.173)	(.170)	(.195)	(.107)
Father's education	Yes	Yes	Yes	Yes	Yes	Yes
Mother's education	Yes	Yes	Yes	Yes	Yes	Yes
State of High School	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4440	3944	3764	3571	3291	4393
R^2	.373	.375	.376	.383	.380	.232

* p<0.10, ** p<0.05

Note: Dependent variable is log monthly earnings in 2003. The first row indicates the selection criteria for the sub-sample included in the estimation. Clustered standard errors in parenthesis.

Table 5: Probit Estimation for Selection into Employment

	1	2	3	4
CenExam*Grade	.0069	.0056	.0087*	.0083
	(.006)	(.006)	(.005)	(.005)
Grade	–.0048**	–.0050**	–.0043**	–.0043**
	(.000)	(.000)	(.000)	(.000)
CenExam (d)	–.0113**	–.0135**	–.0116**	–.0114**
	(.047)	(.042)	(.012)	(.013)
Male-dummy (d)	.0414**	.0403**	.0441**	.0441**
	(.000)	(.000)	(.000)	(.000)
Age	–.000714	–.000687	–.000378	–.000521
	(.434)	(.437)	(.682)	(.566)
Fachabitur (d)			–.0321*	–.0440**
			(.051)	(.043)
Gymnasium (d)				–.0116
				(.309)
Father's education	No	Yes	Yes	Yes
Mother's education	No	Yes	Yes	Yes
Observations	5366	5333	5333	5333
R^2	.010	.013	.015	.015

* p<0.10, ** p<0.05

Note: Dependent variable is an employment dummy taking the value one if the individual reports positive earnings in 2003. Results are reported as marginal effects, dF/dx, evaluated at the sample mean. Marginal effects and standard errors for the interaction terms are based on the consistent estimator for interaction effects in non-linear models presented in Ai and Norton (2003).(d) indicates discrete change of dummy variable from 0 to 1. Clustered standard errors in parenthesis.

Table 6: Evidence on applications, interviews and job offers

	Applications	Interviews	Job offers	Interview/ application	Offers/ application
CenExam	–1.080	–.101	.139*	.0361**	.0647**
	(1.080)	(.106)	(.079)	(.013)	(.024)
Grade	5.958**	.186	.143**	–.0878**	–.0425**
	(.867)	(.118)	(.052)	(.006)	(.016)
Male-dummy	4.254**	.531**	.102	.0211	–.0237
	(.894)	(.143)	(.127)	(.015)	(.029)
Age	.122	–.0590**	.00270	–.000591	–.00228
	(.179)	(.017)	(.019)	(.002)	(.004)
Fachabitur	7.153**	.278	.0121	–.0545**	–.123**
	(1.458)	(.256)	(.246)	(.017)	(.026)
Gymnasium	.221	.217	–.0864	.0327	.0130
	(.778)	(.152)	(.173)	(.023)	(.013)
Constant	–3.381	4.619**	1.835**	.674**	.652**
	(5.945)	(.448)	(.759)	(.071)	(.132)
Father's education	Yes	Yes	Yes	Yes	Yes
Mother's education	Yes	Yes	Yes	Yes	Yes
Observations	4204	3723	3679	3659	3542
R^2	.040	.011	.005	.042	.016

* p<0.10, ** p<0.05

Note: Dependent variables are reported above. All dependent variables are measured in 1998. Clustered standard errors in parenthesis.