

CANDIDATE SELECT GMBH

CASE Study

Efficiency of candidate select at
Simon Kucher & Partners

SIMON ♦ KUCHER & PARTNERS

Strategy & Marketing Consultants



candidate select GmbH

Using data on 180,000 German students as well as 150,000 programme-level grade distributions the candidate select GmbH (hereafter: CASE) developed an algorithm that allows companies to automatically assess their applicants' academic performance.

The CASE Algorithm has multiple stages. First, it compares the applicant to graduates of the same study programme determining his or her position (percentile) in the local grade distribution. Having extracted how good the applicant was relative to fellow students in the same programme we assess how competitive these students have been. In this second step, study programmes are ranked according to the cognitive as well as non-cognitive ability of the students.^{1,2} With the help of these two steps it is possible to understand how good an applicant is compared to his fellow students and how good the fellow students are. To combine the two different parts of information with each other an optimisation algorithm weighs the two scales by defining a similarity score between different applicants. This final step answers the question to what extent an applicant, who is performing well at a school with a lower ranking, can overtake worse performing students at better ranked institutions. The final CASE Score takes all these factors into consideration and allows to compare applicants between different universities, study programmes, degrees and graduation years.

1 Introduction

There are many reasons to believe that a grade is little informative of the labour market success of an applicant. For one, Germany uses an absolute grading system that is interpreted differently by different higher education institutions. Yet also, different study programmes attract students of different background, skill-level and motivation and, thus, unsurprisingly differ in competitiveness and quality. CASE sets out to supply the additional information required to obtain a high-quality signal out of grades.

One way to verify the efficiency of the CASE Score is to analyse the success of the applicants that were assessed by CASE. Since it takes time to be able to review the performance of employees the data basis for such analyses is building up slowly. In this study CASE uses the performance of candidates in the recruiting process as a first indicator of the efficiency of CASE. We are grateful towards our first day to day customer Simon Kucher & Partner (hereafter: SKP) for supplying us with this data and enabling us to conduct this study.

¹Non-cognitive skills are measured with a 50 item Big5 personality battery from the International Personality Item Pool. Cognitive skills are assessed using three different test batteries that we developed: (1) Matrix tests and visual puzzles to assess fluid intelligence and quiz type questions to reveal crystallised intelligence.

²Data has been collected as part of our own Fachkraft study with so far 180,000 German students participating.

2 Study Design

CASE has supplied its customer SKP with its elaborate reports between April and July 2016. During this period all applicants of SKP have been analysed based on the CASE Algorithm. From this group SKP selected a sample of 39 real applicants that are subject of this CASE Study. For this sample the following information are available: (1) Input variables - university, study field, degree, year of graduation and final GPA, (2) CASE assessment - percentile study programme grade distribution and percentile CASE Score and (3) Outcome variables - invitation numerical test dummy, result numerical test and invitation interview dummy. The analyses presented in the next section evaluate to what degree the output variables can be estimated based on input variables as well as CASE assessment scores. The hiring process of SKP is displayed in Figure 1.

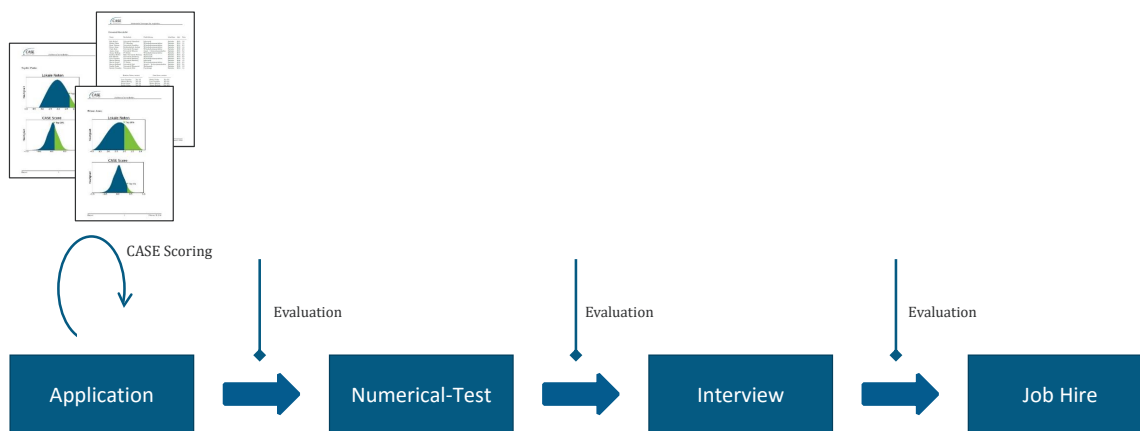


Figure 1: Hiring process of Simon Kucher & Partner.

3 Results

Grades are - per se - an imperfect estimate for job performance. The usage of an absolute grading system, the differences in mean grades and grade distributions, and the variation in quality between study programmes make a useful comparison of grades between different universities, study fields and degrees virtually impossible. The CASE Algorithm, on the contrary, puts grades into perspective through the calculation of CASE Scores. To verify this logic, we estimate the correlations between absolute grades and CASE Scores on the one hand and the outcome variables on the other hand.

The correlations reveal that grades do not possess any significant predictive power on each of the three outcome variables. Noise in grading information is too large to predict recruiting success based on absolute grades. By contrast, CASE Scores add the necessary contextual information to grades. The result is that CASE Scores show moderate to strong correlations with the three outcome variables. These correlations are statistically highly significant, even on the basis of this limited sample.

	Grade	CASE Score
Numeric Dummy	-0.23	-0.52*
Numeric Score	-0.30	-0.50*
Interview Dummy	-0.29	-0.42*

Table 1: Correlation between grades and CASE Scores with a dummy variable for an invitation to a numeric test, the result of the numeric test and a dummy for an invitation to a job interview. The (*) indicates statistical significance at $\alpha = 0.05$.

More emphasis can be put on the relationship between the results of the numerical test with absolute grades and CASE Scores respectively. Figure 2 displays two scatter-plots, each with the numerical test score on the vertical axis. The left panel shows a clutter of points which verifies that grades have indeed little if any predictive power of recruiting success. The right panel displays the significant negative relationship between numerical test results and CASE Scores, i.e. the higher the CASE Score percentile (top 1% being the best), the higher the performance during the numerical test.

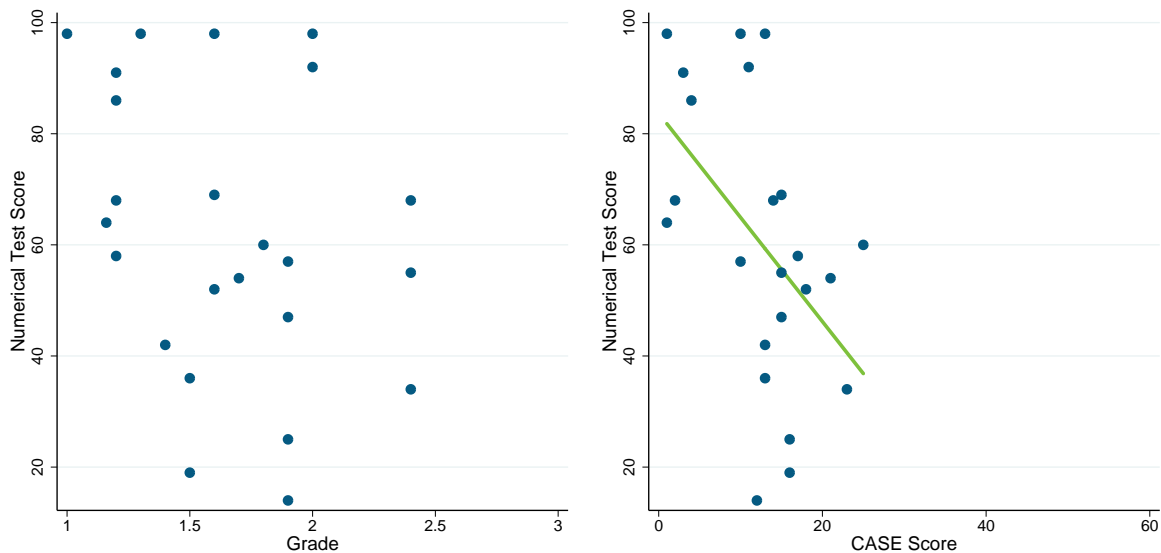


Figure 2: Relationship between numerical test score with absolute grades (left panel) and CASE Score (right panel). Line of best fit included whenever statistically significant at $\alpha = 0.05$.

Moreover, CASE Scores can be used to forecast interview invitations. The correlations in Table 1 already show statistical evidence that this is in fact possible. Estimating a logistic regression model which predicts the chance of an interview based on a CASE Score reconfirms this finding. At a p-value of 0.018 CASE Scores are highly significant. The estimates of this model can be used to display the relationship between CASE Scores and the chance to be invited for an interview. Figure 3 produces this graph. It can be seen that due to the high quality of applicants the probability to be invited for an interview drops to very low levels as soon as a candidate is not part of the top 20% of the German graduate population.

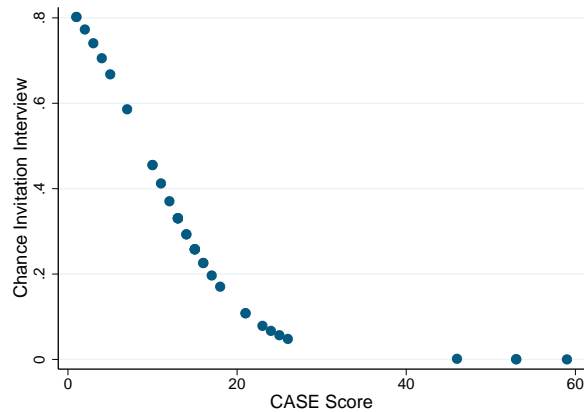


Figure 3: Estimates of a logistic regression model that derives the effect of CASE Scores on Interview Invitations. Relationship is statistically significant at $\alpha = 0.05$.

4 Conclusion

In this short paper we review how well different outcome variables (numerical test results, test and interview invitations) could be predicted by university grades and CASE Scores respectively. We find no meaningful correlation between grades and any of the outcome variables. By contrast, CASE Scores show moderate to strong correlations with the three outcome variables in the predicted direction. While the sample size with 39 applicants is small it is noteworthy that the correlations found are all statistically significant at $\alpha = 0.05$.

Authors and Contact

candidate select GmbH
Dr. J.N. Bergerhoff, Dr. P.K. Seegers, M. Hoyer
Email: info@candidate-select.de
Web: www.candidate-select.de
