

Siegenthaler, Michael

## **Can a standardised aptitude test predict the training success of apprentices? Evidence from a case study in Switzerland**

*Empirical research in vocational education and training* 3 (2011) 2, S. 105-128

urn:nbn:de:0111-opus-82553



in Kooperation mit / in cooperation with:



<http://www.skbf-csre.ch>

### **Nutzungsbedingungen / conditions of use**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.  
By using this particular document, you accept the above-stated conditions of use.

### **Kontakt / Contact:**

**peDOCS**  
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)  
Mitglied der Leibniz-Gemeinschaft  
Informationszentrum (IZ) Bildung  
Schloßstr. 29, D-60486 Frankfurt am Main  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

# Can a standardised aptitude test predict the training success of apprentices? Evidence from a case study in Switzerland

Michael Siegenthaler\*

*Swiss Economic Institute (KOF), Zurich*

## Abstract

Due to widely spread distrust in the signalling value of school grades, Swiss employers require external standardised aptitude test results in the process of recruiting new apprentices. However, the predictive quality of such tests has never been thoroughly researched. Therefore, this case study analyses whether aptitude tests can improve the ability of firms to predict success in apprenticeship training. I find that such tests do not add information that explains the success in VET schooling (school grades during the first and second years of apprenticeship training), the probability of unexcused vocational school absences, or the likelihood of a premature termination of the apprenticeship contract.

*Keywords: apprentice hiring, aptitude test, predictive validity, screening*  
*JEL classifications: I21, J24, M51, M53*

## 1. Introduction

During the last fifteen years, standardised aptitude tests for apprenticeships provided by private firms have increasingly influenced the transition of adolescents from lower secondary schools into dual training programmes in Switzerland. Many enterprises that hire apprentices currently require such a test from their applicants, particularly in the commercial sector and the retail business. This study examines the predictive power of one such test: the «Multicheck retail sale». Is the test a valid predictor of the vocational success of apprentices? By answering this question, we also gain insight regarding whether the test can improve the apprentice selection of a firm.

The success of aptitude tests for apprenticeships is closely linked to the deficiencies of performance measures provided by lower secondary schools (Imdorf, 2009;

\* Address for correspondence: Swiss Economic Institute (KOF), Weinbergstrasse 35, CH-8092 Zurich. E-mail: siegenthaler@kof.ethz.ch. Phone: +41 44 633 93 67. I would like to thank Stefan Wolter for his assistance and his comments during the entire research process. I would like to thank Ralph Hardegger, Michèle Oswald, Bernd Schauer, Simon Siegenthaler, participants of the Brown Bag Seminar at ETH Zurich, participants of the VET congress in Zollikofen, and an anonymous referee for helpful comments and suggestions.

Moser, 2004). Employers require the tests because they cannot infer the true scholastic abilities and cognitive skills of their candidates on the basis of school grades and the types of schools that applicants attended. Employers' mistrust in school grades may be justified. School achievements represent a noisy and biased measure of the effective scholastic abilities of pupils; they are only weakly correlated with «objective» measures of scholastic and cognitive ability, such as PISA scores or IQ tests (Baron-Boldt, 1989; Bauer & Sheldon, 2008; Bishop, 1994; De Paola & Scoppa, 2010; Kronig, 2007; Lindahl, 2007). Moreover, the Swiss school system is very opaque. For instance, Swiss cantons and sometimes even municipalities differ considerably in the amount of educational tracking and curricula. School reforms have increased the heterogeneity among school-leavers. Thus, comparing the academic achievements of applicants for an apprenticeship position can entail substantial problems for a firm.

The tests seem to alleviate this informational deficit. The standardisation of these tests enables the comparison of the scholastic and cognitive abilities of candidates in a simple and economic way. Additionally, the tests economise the candidate selection process. Hence, it is not surprising that studies have found that these test results can considerably influence an applicant's chances of being employed as an apprentice (Moser, 2004; Mühlemann et al., 2007).

By improving the «signalling» of the scholastic and cognitive competences of pupils to employers (Bishop, 1994), the tests could improve the matching between applicants and firms and hence increase social welfare (Costrell, 1994). However, can the tests keep their promises? What is their effective value as predictors of vocational success? Few authors have attempted to answer these questions. To my knowledge, the only scientific attempt to examine the validity of Swiss aptitude tests is a study by Widmer (2006), who examined a test for the commercial sector.<sup>1</sup> This lack of scientific attention is surprising given that the tests can have a substantial influence on the educational careers of young adults.

This paper contributes to filling this research gap by assessing the predictive power of the «Multicheck retail sale» (Multicheck Detailhandel und Service). All major Swiss enterprises in the retail business require apprenticeship applicants to submit their results from this test. The informational value of the test is examined using a self-collected data set containing information regarding the personal statuses, lower secondary school reports, Multicheck results, and vocational school reports of 334 apprentices employed at Migros, the largest employer in the retail business in Switzerland. The main outcome variable considered is vocational school grades in the first and third semesters. The results are verified using the probability of unexcused vocational school absences and of premature apprenticeship terminations as additional criteria of vocational success.

<sup>1</sup> However, the study is partly limited by small sample sizes. The investigations of Multicheck (2010) regarding the validity of their tests do not meet scientific requirements because they lack documentation.

The regressions demonstrate that the overall Multicheck score (Gesamtergebnis) is a poor predictor of vocational school performance in both semesters considered. School grades and the types of schools the applicants attended are noticeably more reliable sources of information for recruiters. These findings are robust to different specifications and estimation methods. Additionally, they carry over to the other outcome variables considered.

The structure of the paper is as follows. In the next section, I discuss the informational value that can be theoretically expected from standardised aptitude tests for apprenticeships. Section 3 provides information regarding the data and the sample. The following sections present the main regression results (Section 4) and their robustness (Section 5). The paper concludes after the discussion of the two additional outcome variables (Section 6).

## **2. Multicheck and the prediction of training success**

### ***2.1 The apprenticeship system in Switzerland***

In Switzerland, dual apprenticeship training is the predominant form of upper secondary education. Each year, approximately sixty percent of a cohort of school-leavers who pursue some type of post-compulsory schooling enrol in an apprenticeship training programme. Swiss adolescents can choose among approximately 230 different training occupations. Once an occupation is chosen, the adolescents apply to firms that are recruiting new apprentices. Successful candidates then receive an apprenticeship contract that, among other things, determines the quantity and quality of the training provided by the firm and the apprentice pay.

The duration of the apprenticeships is two, three, or four years, depending on the specific occupation. During the training, the apprentices are provided with theoretical education and practical training. The theoretical education is primarily provided by state-organised and publicly financed vocational schools. The apprentices spend between one and two days per week in these schools and obtain general education and theoretical knowledge pertaining to occupation-related issues. Additionally, an apprentice receives further formal training at the workplace throughout the training programme. During the training, apprentices regularly write examinations and receive school records each semester, and their performance at the workplace is assessed by the apprenticeship trainers of the firms. Furthermore, an extensive final exam is administered at the end of the programme. Each successful graduate receives a national certificate. This certificate attests that he or she has gained the professional qualification to perform skilled work in the respective occupation (for more information on the Swiss apprenticeship system, cf. Mühlemann et al., 2007, and Mühlemann et al., 2009).

## 2.2 *The Multicheck retail sale*

In Switzerland, aptitude tests for apprenticeships are provided and administered by specialised private firms. «Multicheck» and «basic-check» are the most widely used tests. Even small and medium-sized enterprises require the results of these tests from their applicants. By its own account, the firm Multicheck is the largest provider of standardised aptitude tests in Switzerland. More than 30,000 Swiss adolescents take one of the six different «Multicheck Junior» tests each year. Multicheck refers to its tests as «job-relevant aptitude assessments» on its homepage. According to Widmer (2006, p. 49), the aim of these tests is «to predict the practical and especially theoretical success of the vocational training of apprentices».

The specific test version of Multicheck studied in this paper is the «Multicheck retail sale» (Multicheck Detailhandel and Service). This test is currently mandatory for nearly all adolescents applying for apprenticeships in the retail business. Evaluations of the test can be obtained for different apprenticeships in the field. This study concentrates on the version for potential «retail sales specialists» (Detailhandelsfachmann/-fachfrau). Approximately 10,300 adolescents completed this version of the Multicheck test in 2009. This test was considered in this investigation primarily because of the quantitative importance of the test.

The test assesses knowledge in school subjects (German, French, English, or both foreign languages, depending on the choice of the candidate, and numeracy) and general cognitive skills (logic, concentration, and retentiveness). The results are presented separately for each of the six (or seven) test sections. In addition, an overall score (Gesamtergebnis) is generated. The scale used to present the scores ranges from zero to hundred. This scale is relative: an overall score of fifty points indicates that fifty percent of previous test takers have received higher scores and fifty percent have received lower scores. The relative scale does not seem to allow statements regarding a test taker's aptitude to begin a certain apprenticeship. However, the evaluation of Multicheck contains such an indication: overall test scores of fewer than forty points are «insufficient», scores between forty and sixty points are «sufficient», and test takers with more than sixty points «exceed» the requirements of an apprenticeship (cf. Appendix A).<sup>2</sup>

## 2.3 *Assessing the predictive power of the test*

Firms use the test to identify the quality of their candidates. Thus, this test is a «screening device» (Arrow, 1973; Stiglitz, 1975).<sup>3</sup> If we determine whether the test

2 Although Multicheck states that the decision to accept a candidate is left to the company, it nevertheless refers to «experience» that «has shown that candidates with an overall average below forty percent do not complete their apprenticeship, or only with struggle» to justify the thresholds. The argument is not very convincing: How can Multicheck ensure that exactly 39.5 percent of the test takers do not fulfil or «insufficiently» fulfil the requirements of a particular apprenticeship?

3 While firms screen, applicants signal. However, the situation analysed in this paper is less closely related to the concept of signalling (see Spence, 1973; Weiss, 1995) than to screening. Signalling refers

is a valid predictor of the training success of apprentices, we may gain information regarding the test's «usefulness» as a screening device. Its usefulness is given by the amount by which the test improves a firm's ability to assess an applicant's chances of vocational success. In this respect, the «informational value» of the test depends on the informational surplus that it adds to the selection decision of a firm. If all of the information provided by the test is already inferable from other sources (for instance, from school grades in the application dossier), then the test would not be valuable for employers. Thus, the informational value of the test should be analysed conditional on other information available in the selection.<sup>4</sup>

The validity of the multivariate regression approach depends on the choice of the control variables. In particular, the choice is constrained by the recruitment practices at Migros. This firm uses Multicheck nearly exclusively for the pre-selection of applicants (similar to the firms examined by Moser, 2004). The test has no major role in the later stages of the recruitment process. Hence, if we included explanatory variables that actually explain vocational training success but are unavailable to the firm at the time of the pre-selection decision, then we may obtain biased results pertaining to the informational value of the Multicheck test for the firm.

Consequently, the potential right-hand variables in the regressions should be derived from the application dossier. The most important sources of information therein are lower secondary school records. They contain information regarding school grades and the types of schools that the applicants attended. Survey evidence and empirical studies document that school grades and types of schools have a major effect on the probability that an applicant will be employed (Häberlin et al., 2004; Imdorf, 2009; Moser, 2004; Stalder et al., 2008). A second source of information in the application dossier is the unalterable personal attributes of a candidate. Attributes that are likely to influence the apprentice selection of firms are gender and social and cultural background, such as nationality, native language, and regional origin (Amos et al., 2009; Bertschy et al., 2008; Häberlin et al., 2004; Imdorf, 2005; Moser, 2004; Stalder et al., 2008).

What variable should we use on the left-hand side of the regression, i.e., which operationalisation of vocational success is relevant in the sense that Multicheck attempts to predict this success? There are at least three reasons why the most relevant outcome variable in our context is vocational school GPA. Firstly, the test mainly aims to forecast theoretical success in vocational training (see Section 2.2). Secondly, Migros uses the test nearly exclusively to assess an applicant's chances of meeting the academic requirements of an apprenticeship. Finally, problems in voca-

to a self-selective mechanism that induces job searchers to gather a costly «signal» of human capital - a mechanism underlying school grades. Multicheck is more closely related to screening (i.e., a firm's assessment of applicants).

4 The chosen approach corresponds to the psychological concept of «incremental validity»: Does Multicheck add to the prediction of a certain criterion above that which can be predicted by other sources of data (Hunsley & Meyer, 2003)? This approach differs from the approach utilised in the studies of Widmer (2006) and Multicheck (2010), in which the «predictive validity» of the test is analysed (i.e., simple correlations between a test score and a criterion of vocational success).

tional school are the most important causes of premature apprenticeship terminations (Stalder & Schmid, 2006).

Nevertheless, I control the robustness of the results using two additional outcome variables: the probability of unexcused vocational school absences and the probability of premature apprenticeship termination. Both situations are costly for the firm<sup>5</sup>, are interrelated to vocational school grades and the respective other outcome variable, and have a relatively close relationship with the practical performance of apprentices (Imdorf, 2009; Stamm et al., 2010).

The reasoning of the previous paragraphs boils down to the regression models presented in Section 4. What hypothesis can we propose regarding the informational value of the Multicheck retail sale in these regressions? First, the test should benefit from its standardisation, as this standardisation prevents errors due to relative measurement, and these errors are important sources of noise in school grades. Secondly, the test measures the academic and scholastic abilities of adolescents. Studies have shown that such tests can be valid predictors of labour market and educational success.<sup>6</sup> Moreover, if we consider that school grades and school levels are likely to imperfectly represent the actual cognitive or scholastic abilities of pupils, Multicheck may significantly contribute to the prediction of vocational success (i.e., its incremental validity could be considerable).

However, three factors could limit the informational value of the test. First, the test assesses the skills of test takers during a period of approximately two and a half hours. In contrast with school grades that often represent an average measure of several assessments, this relatively short one-time testing method may be noisy (De Paola & Scoppa, 2010; Grant, 2007). Secondly, Stamm et al. (2010) found that more intellectually gifted apprentices do not necessarily perform better. If these results are relevant to the context analysed in this paper, the cognitive part of the Multicheck test may be a poor predictor of success in vocational training. Thirdly, shortcomings in the construction of the test could undermine its informational value. In this respect, Moser (2004) has expressed doubts regarding the theoretical quality of the aptitude tests employed in Switzerland, especially concerning the fairness and internal coherence of these tests.

### 3. Data

The data set covers apprentices from three (of ten) cooperatives of Migros, the largest employer in the Swiss retail business with a market share of 36.8 percent in 2009.

<sup>5</sup> The former situation is costly because the vocational instructors at Migros will have to discuss the reasons for the absence with the apprentice, and the latter situation is costly because premature apprenticeship terminations (as with any turnover) are costly in terms of time and financial outlay (Mühlemann et al., 2007).

<sup>6</sup> For tests of academic achievements, such as PISA or the ACT, see, for example, Bertschy et al. (2008), Bettinger et al. (2011), or Stalder et al. (2008). For tests of cognitive skills, such as IQ tests, cf. Grant (2007), Heckman & Vytlačil (2001), or Murnane et al. (2001).

The cooperatives considered are Migros Aare, Lucerne, and Zurich. The sample covers all apprentices that began their apprenticeship as «retail sales specialist» in August 2007 or August 2008.<sup>1</sup> I gathered the data directly from the personnel files of each cooperative. The sample consists of 334 apprentices: 142 apprentices from Migros Aare, 84 apprentices from Migros Lucerne and 108 apprentices from Migros Zurich. Coding issues and the definitions of the variables gathered are discussed in Appendix B.

The vocational school grades collected for this study stem from the first and third semesters of vocational school. Thus, because the duration of the retail sales specialist apprenticeship is six semesters, the grades considered in this paper represent training success in an early phase of the apprenticeship. To study grades from an early phase of the apprenticeship rather than grades from the final exam is not a major drawback. First, grades from the early semesters are relatively closely correlated with the grades received for the final exam of the apprenticeship. Secondly, the use of early grades ensures greater external validity of the study because the specific firm cannot yet strongly influence them.

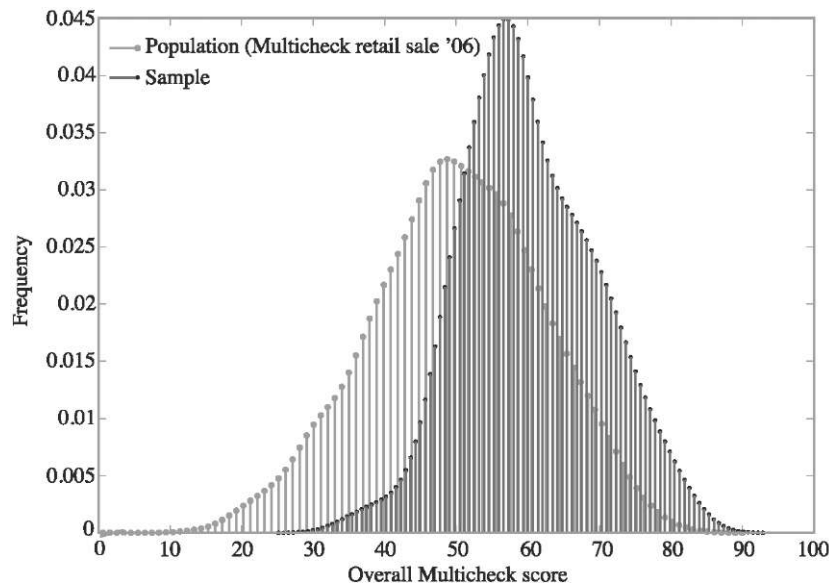
An additional contributor to a greater external validity of the study is that the vocational training and learning situations of apprentices vary across and within each cooperative. First, the hiring and management of apprentices differ between the three cooperatives, and secondly, apprentices are largely trained on the job by an apprenticeship trainer at the branch offices where they work.

However, the study is partly limited by a selection problem: the sample has too few «insufficient» overall Multicheck scores. In particular, the sample probability that an applicant with an insufficient test score is employed at one of the three Migros cooperatives is only 1.2 percent. This pattern arises because the Migros cooperatives assign considerable weight to the Multicheck score in the selection of apprentices and because adolescents with low test scores may stop searching for retail sales specialist apprenticeships. Figure 1 illustrates the problem by comparing the sample distribution of the overall Multicheck scores (Gesamtresultat) with the distribution of overall scores from the population of test takers.<sup>7</sup> As a result of the compressed sample distribution of the overall Multicheck scores, I can examine the informational value of only the upper two-thirds of the score. I will discuss the implications of the sample selection problem in the robustness section.

7 The latter data were provided by Multicheck. The firm sent raw data of the test results of all Multicheck 06/07 (2006 version of the test) test takers. Multicheck did not provide the actual overall scores but only the results in each individual test section. I generated an overall test score from the results of the subsections under the condition that the overall score must have a mean of 50 points according to its scaling (cf. Section 2.2). The resulting figure was approved by Multicheck. The histograms were computed using a Gaussian kernel.



**Figure 1. Sample distribution of overall Multicheck scores (dark) and (estimated) population distribution of overall scores (light)**



#### 4. Predicting success in VET schooling

Table 1 shows the cross-correlations of the five most important variables. Evidently, the overall Multicheck test scores do not (even negatively) correlate with school grades from lower secondary school. This result is consistent with the findings of Widmer (2006) for the «Multicheck commercial sector». The correlations between lower secondary school grades and GPAs in vocational school meet the expectations.

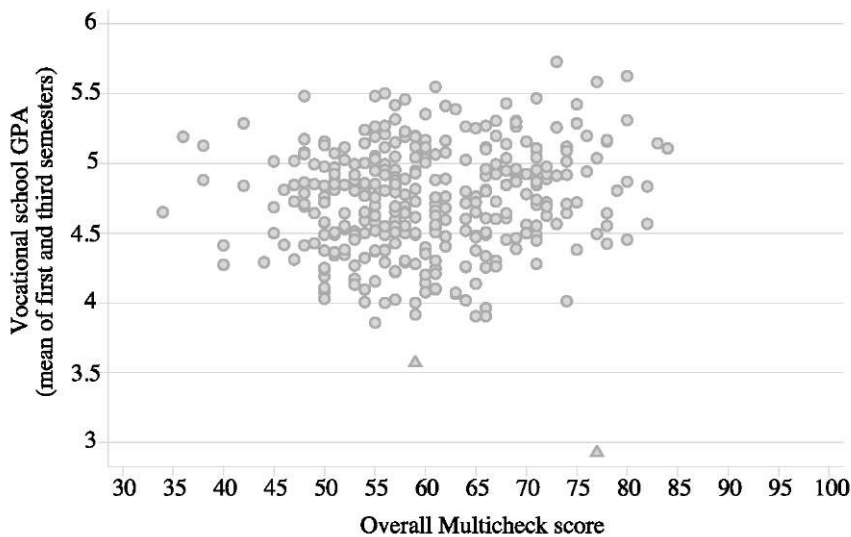
However, the «predictive validities» of 0.113 and 0.12 of the overall Multicheck scores do not meet the expectations. They are relatively low compared with the correlations reported by Widmer (2006), and they are substantially below the correlation of 0.4 established in the Multicheck (2010) study that considered the same Multicheck test version. In addition, the predictive validity of the overall score is weaker than those commonly established for tests of general cognitive ability. The correlations usually lie between 0.3 and 0.5 (see, e.g., Schmidt and Hunter, 1998).

**Table 1: Correlations between lower secondary school GPAs, overall Multicheck scores and vocational school GPAs**

	(1)	(2)	(3)	(4)	(5)
GPA of newest lower secondary school record (1)	<b>1</b> n=334				
GPA of second-newest lower secondary school record (2)	<b>0.812**</b> n=326	<b>1</b> n=326			
Overall Multicheck score (3)	<b>-0.011</b> n=334	<b>-0.094*</b> n=326	<b>1</b> n=334		
GPA in vocational school, first semester (4)	<b>0.275**</b> n=331	<b>0.22**</b> n=323	<b>0.113**</b> n=331	<b>1</b> n=331	
GPA in vocational school, third semester (5)	<b>0.291**</b> n=324	<b>0.223**</b> n=316	<b>0.119**</b> n=324	<b>0.737**</b> n=323	<b>1</b> n=324
<b>** p &lt; 0.01, * p &lt; 0.05</b>					

The small correlation between overall Multicheck scores and average vocational school grades is illustrated in Figure 2. The grading scale in Switzerland ranges from the «excellent» grade of 6.0 to 1.0 with 4.0 as the lowest sufficient grade. The scatter plot shows a considerable dispersion of values.

**Figure 2. Overall scores in the Multicheck test and average grades in vocational school (Triangles: Two outliers subsequently excluded from the analysis)**



Furthermore, it demonstrates that a test taker who «exceeds» the requirements of the retail sales apprenticeship according to the test (i.e. a test taker who scored above sixty points in the overall Multicheck score) will not necessarily perform strongly in vocational school. Similarly, some apprentices with relatively poor Multicheck results (between forty and sixty points in their overall scores) performed well in vocational school. When examining the figure, one may observe that the Multicheck test cannot assure that adolescents with «insufficient» test scores (i.e. Multicheck results of fewer than forty points) will underperform in vocational school.

However, as discussed in Section 2, this paper aims to establish the predictive power of Multicheck in a multivariate regression model. By scrutinising the data, one finds that the outcome variable - vocational school grades - is quite strongly clustered. In particular, the levels and, to a lesser extent, the variances of vocational school grades are different across the 23 vocational schools of the sample despite the equal performance standards among all vocational schools. An elegant approach to handling the problem of clustering in the outcome variable is the use of a nonlinear probability model with an appropriate outcome variable (cf. Fielding et al., 2003). Therefore, I construct a grade variable that ranks an apprentice relative to the performance of apprentices attending the same vocational school. More precisely, the new outcome variable (named *Grader*) takes on the value zero (low grader) if the average grade of an apprentice belongs to the lowest quartile of the sample distribution of grades in the vocational school that she or he attends. This value is two if she or he belongs to the quartile of top graders of his school; otherwise, the value is one (medium grader).

One may infer that the procedure amounts to assuming that the observed differences in vocational school grades are only due to the varying degrees of difficulties in the vocational schools or, in other words, to assuming that the apprentices in different vocational schools are equally able. First, this assumption would not be restrictive, as the selection into the different vocational schools is largely exogenous and the selection of Migros should ensure the relative homogeneity of the sample.<sup>8</sup> Secondly, the variable is a meaningful operationalisation of success in vocational training even if this exogeneity is not given: it is valuable for the employer to know whether an apprentice is likely to belong to the group of top performers or the group of underperformers in a specific vocational school.

The transformation of vocational school GPAs in a categorical variable has a further favourable effect in our context: I assume that vocational school grades are hierarchical; I do not assume that they have a continuous interval scale (as is necessary in a linear regression model with GPA as a dependent variable). Due to the lack of an interval scale underlying school grades, several authors have argued that grades

<sup>8</sup> The selection of different vocational schools is determined by the branch office of the apprentice. As such, it is exogenous to the apprentice. Furthermore, because the recruiters at Migros were not aware of the different grading practices at vocational schools and because the placement into branch offices is largely determined by the domicile of the apprentice and job openings at branch offices, the selection also seems exogenous to the firm.

should be analysed using an ordered probit or ordered logit model (Fielding et al., 2003; Grant, 2007; Sund, 2009).

To identify our model, we must specify a latent variable. Specifically, to relate the three outcomes of the observed categorical with the latent variable, we must estimate two threshold values ( $\alpha_1$  and  $\alpha_2$ ):

$$Grader_i = \begin{cases} 0 & \text{if } -\infty < Grader_i^* < \alpha_1 \\ 1 & \text{if } \alpha_1 < Grader_i^* < \alpha_2 \\ 2 & \text{if } \alpha_2 < Grader_i^* < \infty \end{cases} \quad (1)$$

In this formula,  $Grader_i^*$  denotes the underlying latent variable. We estimate the following linear model:

$$Grader_i^* = \beta_1 + \beta_2 MCoverall_i + \beta_3 GPA_i + \beta Z_i + \varepsilon_i \quad (2)$$

In (2),  $MCoverall_i$  represents the overall score of the Multicheck retail sale test,  $GPA_i$  is the average school grade of the two lower secondary school records collected,  $Z_i$  is a vector of control variables, and  $\varepsilon_i$  is the residual of the regression. In a reduced model, the vector of exogenous control variables,  $Z_i$ , comprises (dummy) variables for age, gender, and native language; a variable indicating whether an applicant attended a lower secondary school in an urban region or a rural region; variables representing the different types of schools; and a dummy variable indicating whether an apprentice attended a course to bridge gaps in training (CBGT). The full model adds eleven cantonal dummy variables (with the canton of Aargau as the reference group) and two dummy variables for the Migros cooperatives (with Migros Aare as the reference group) to the vector of controls (cf. Appendix B for the specification of these controls). Because this paper aims to estimate a consistent coefficient for the Multicheck result, some variables are included in the regressions although they are insignificant.

The model in (2) is identified by assuming a distribution for the residual  $\varepsilon_i$ . The most popular choices are the logistic and standard normal distributions that lead to the ordered logit or ordered probit model, respectively. The test regressions show that the differences between the two models are minor. I present the results for the ordered probit model.

The ordered probit estimates of equation (2) are presented in Table 2.<sup>9</sup> The overall score of the Multicheck test is very inconsistently (if at all) useful as a predictor of average vocational school grades in the first and third semesters. However, the average lower secondary school grade is a valid predictor of theoretical training success. Moreover, the signalling value of secondary school *records* profits from the predic-

<sup>9</sup> Two outliers are consequently omitted from the regressions of first-semester vocational school GPAs because they have substantial leverage related to the coefficient of the overall Multicheck score. This leverage is illustrated in Figure 2, in which the two outliers are highlighted as triangles.

tive power of the type of school coefficients. Apprentices that attended an intermediate- or advanced-level school perform significantly better than those from the basic-level schools.

Table 2: Determinants of vocational school GPAs in the first and third semesters (ordered probit regression)

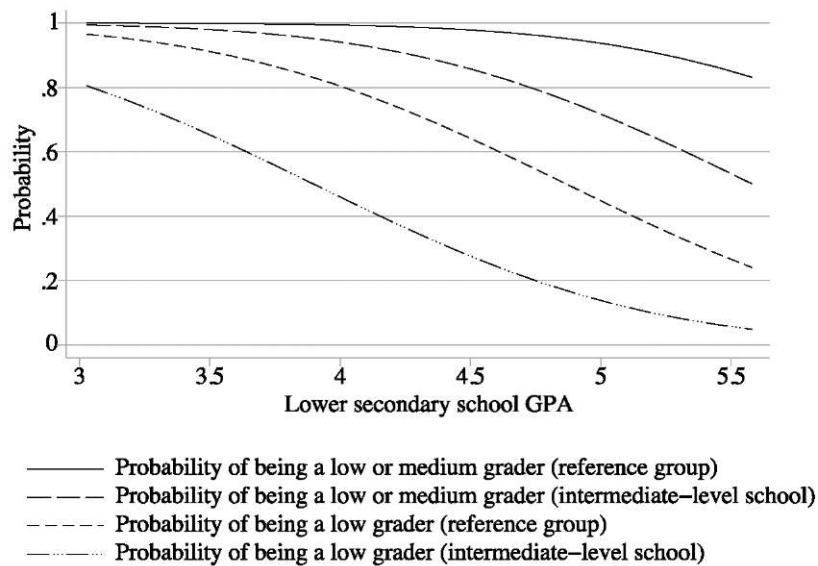
Dependent variable:	Vocational school GPA 1 (Grader* 1)		Vocational school GPA 3 (Grader* 3)	
	(1)	(2)	(3)	(4)
Overall score MC	0.009 (0.007)	0.005 (0.008)	0.013+ (0.007)	0.010 (0.008)
GPA	0.644** (0.158)	0.991** (0.183)	0.487** (0.158)	0.730** (0.179)
>17 2/3 yrs	0.169 (0.184)	0.124 (0.189)	-0.126 (0.186)	-0.161 (0.190)
<16 yrs	-0.079 (0.257)	-0.092 (0.267)	-0.043 (0.257)	-3.88e-05 (0.267)
Female	-0.0203 (0.135)	-0.0583 (0.143)	0.086 (0.136)	0.092 (0.143)
German	0.298* (0.147)	0.389* (0.153)	0.164 (0.150)	0.209 (0.154)
Urban	-0.018 (0.165)	-0.189 (0.173)	-0.064 (0.166)	-0.164 (0.174)
CBGT	0.097 (0.148)	0.129 (0.154)	0.071 (0.148)	0.086 (0.154)
Intermediate-level school	0.558** (0.143)	0.961** (0.174)	0.410** (0.143)	0.744** (0.172)
Advanced-level school	1.235** (0.451)	2.014** (0.500)	1.564** (0.480)	2.140** (0.512)
Cantonal dummies	No	Yes	No	Yes
Dummies for cooperatives	No	Yes	No	Yes
Threshold 1	3.172**	5.119**	2.573**	3.863**
Threshold 2	4.741**	6.787**	4.212**	5.572**
Observations	329	329	324	324
McFadden's Pseudo-R <sup>2</sup>	0.069	0.112	0.056	0.088

\*\* p < 0.01, \* p < 0.05, + p < 0.1

Notes: Standard errors in parentheses

Figure 3 exemplifies the effects of the school variables. This figure shows the probabilities that a hypothetical person, belonging to the reference group and applying with a Multicheck score of sixty points, will be a low grader or a low or medium grader. Figure 3 also contains the outcome probabilities for a person who did not attend a basic-level school (as the person from the reference group) but who attended an intermediate-level school, with everything else held constant. Thus, for instance, the probability that an adolescent applying with a GPA of 5.5 from an intermediate-level school will be a low grader in the first semester of vocational school is below ten percent.

Figure 3. Estimated probabilities that a person from the reference group or from an intermediate-level school belongs to the group of low graders or to the group of low or medium graders



Notes: The reference group of the full model (columns 2 and 4 in Table 2) used for this figure are male and medium-aged apprentices of Migros Aare that applied from a (rural) basic-level school of the canton of Aargau, who do not speak German as a native language, and who did not attend a course to bridge gaps in training.

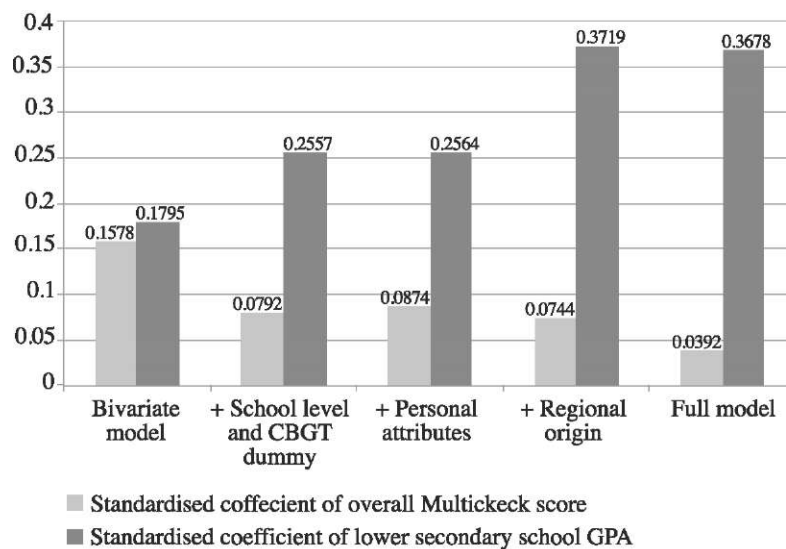
An interesting exercise is whether the low informational value of the overall Multicheck score is the result of the inclusion of other variables. Figure 4 provides the answer to this question using the ordered probit model for first semester vocational school GPAs. The figure compares the (standardised) coefficients of the overall Multicheck scores with those of the average lower secondary school grades when other explanatory variables are added to the model, beginning with a bivariate model. The figure illustrates the appropriateness of the multivariate approach. It shows how cor-

relations would be misleading regarding the informational surplus that the test adds to the selection decision of the firm.

Particularly, the overall Multicheck scores and GPAs have nearly the same informational value for employers if they are considered independently of any other information. However, Multicheck loses much of its informational value once an employer knows whether the Multicheck result is from an attendee of a basic, intermediate, or advanced-level school and whether the adolescent attended a CBGT (second column). This result occurs because attendees of higher-level schools or CBGTs have higher Multicheck scores (as OLS regressions of the overall Multicheck score show). Thus, the Multicheck results mirror the higher average skills of pupils of higher school types, but the «type of school» signal is a stronger predictor of vocational school grades than the Multicheck test result. The predictive power of the test is further reduced if we include lower secondary school grades in the model (last column). In conclusion, the informational surplus that the test adds to the selection of retail sales specialists becomes minor once a recruiter considers (and correctly interprets) information from the elementary school records of an applicant. This finding is contrary to the hypothesis formulated in Section 2.3.

Interestingly, the pattern is reversed if we examine lower secondary school GPAs: the signalling value of school grades increases with the amount of information regarding their context. In particular, employers can interpret school grades more accurately if they consider the school levels and regional origins of the adolescents.

Figure 4: Standardised coefficients of lower secondary school GPAs and overall Multicheck scores when regressors are added to an ordered probit regression



### 5. Robustness

The Multicheck test is a poor predictor of the training success of apprentices at Migros. This result is robust for the following reasons. First, the result appears in all specifications of the ordered probit model. Secondly, the result is confirmed in a school-fixed-effects model of vocational school grades (that is, in a linear regression model that accounts for the differences in mean grades across vocational schools). Finally, the result is not affected by the inclusion of the variable of social background - a variable of the skill of an apprentice's parents (cf. Appendix B).

However, two remarks regarding this finding must be offered. The first remark is demonstrated in Figure 2: the correlation between overall Multicheck scores and vocational school grades is greater for Multicheck results above the sample mean (60.17 points) than for those below the mean:  $\beta_{MCoverall} \leq 60 = 0.052$  and  $\beta_{MCoverall} > 60 = 0.281$ . This result is also confirmed in ordered probit regressions: above-average Multicheck results are slightly more valid in predicting vocational school grades than below-average scores.

The non-constancy of the Multicheck coefficient indicates a limitation of this study that is probably related to the sample selection problem discussed in Section 3. For example, consider the following selection effect: adolescents who received apprenticeships at Migros despite a bad Multicheck score of only forty points may have compensated for their poor test scores by writing application letters that were better than those written by those applicants with the same Multicheck results who were not recruited. If the necessity to compensate for a poor Multicheck score in the pre-selection process of Migros decreases with a better test result, then the selection effect would induce an omitted variable bias: the regressions would lack an unobservable quality of the applicant that is negatively correlated with the Multicheck score. Finally, because the unobservable quality is positively correlated with vocational school grades, the estimate of the Multicheck coefficient would be biased downwards. In other words, we would underestimate the slope of a regression line in Figure 2 because we lack bad graders for low overall Multicheck scores (observations in the lower-left part of the figure) relatively more than we lack bad graders for good overall scores (observations in the lower-right part of the figure).

If such a selection bias occurred, then we would be unable to judge the «usefulness» of the Multicheck test in the selection of Migros (i.e. its value as a predictor of the training success of applicants). We would be limited to drawing conclusions regarding the test's predictive power of the training success of actual apprentices. The quantitative importance of the selection effect on the Multicheck coefficient is difficult to assess.<sup>10</sup> We can attempt such an assessment by examining the size of the Multicheck coefficient for the high test results. Because the importance of the Multicheck results in the pre-selection phase decreases once the candidates «exceed» the requirements of an apprenticeship, the unobservable qualities that are important for

<sup>10</sup> The estimation of a selection model that incorporates the selection decision of the firm was not feasible (due to privacy concerns and time constraints).



the firm's pre-selection are unlikely to be strongly correlated with high test scores. As the point estimate of the Multicheck coefficient is also small for test results above sixty points, we can be rather certain that the sample selection bias of the Multicheck coefficient is small. In addition, lower secondary school grades are also exposed to the selection problem. When comparing the coefficient of school grades with the coefficient of the Multicheck scores, the relative bias should be rather small. In addition, two findings provide strong evidence that the Multicheck test suffers from conceptual problems. First, regressions indicate that several subsections of the Multicheck test function reasonably well in predicting vocational school grades. In particular, the language sections of the test are good predictors of corresponding vocational school grades. For example, the correlation between French scores and vocational school grades in French is 0.412. Second, the predictive power of the test could be improved by a simple measure: generating an own overall score. This is because the Multicheck scores for English and retentiveness are even better predictors of vocational school GPAs than overall test scores, while the 'logic' and 'concentration' sections do not add to and may even undermine the predictive power of the test. The test may fail to accurately measure these cognitive skills, or cognitive skills may not be decisive for success in vocational training (cf. Stamm et al., 2010). Either way, dropping the two test sections would significantly increase the predictive power of the test. <sup>11</sup>

## **6. Other variables of success in vocational training**

Even if vocational school grades are the most important determinants of theoretical success in vocational training, they only partially reflect the aptitude of an adolescent to successfully complete an apprenticeship as a retail sales specialist. Is the Multicheck retail sale a better predictor of other criteria of vocational success? Two further outcome variables are analysed to answer this question.

The first variable is the probability of unexcused vocational school absences that are used as a proxy for the social competence of an apprentice. Unexcused absences in vocational school often express a student's lack of motivation, problems in school, or a missing sense of duty (Lounsbury et al., 2004). Not surprisingly, they are major causes of premature terminations of apprenticeship contracts (Moser et al., 2008).

Because non-attendance can result from problems in school, one might expect that school grades and Multicheck scores can act as a significant predictor of this outcome variable. Among the sample, 16.55 percent or 48 out of 290 apprentices had at least one unexcused absence during the first two years of the apprenticeship (cf. Appendix B for a discussion of the reasons for the reduction in sample size).

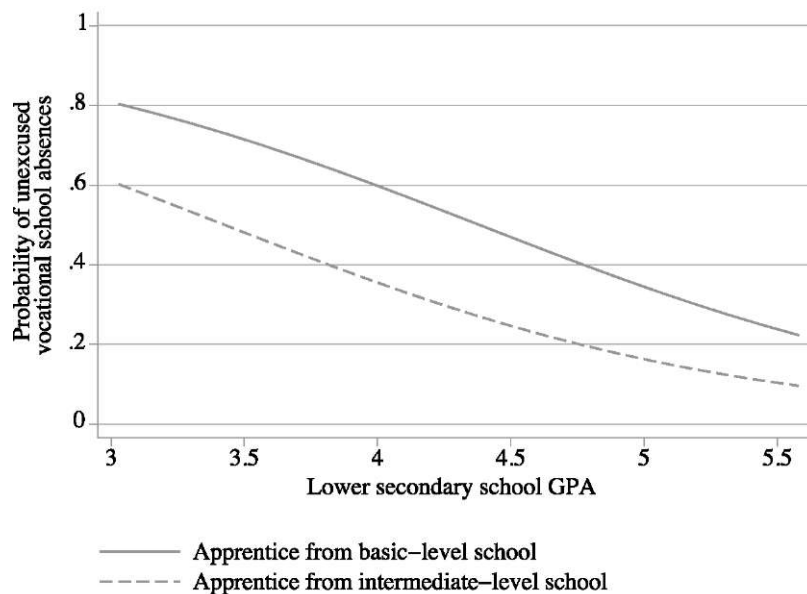
The results of the logit regressions of the binary outcome variable on the set of control variables already used in the models of Section 4 are presented in Appendix C.<sup>12</sup>

<sup>11</sup> Bettinger et al. (2011) found similar patterns when examining the ACT exam that is used in College admission in the US.

<sup>12</sup> The logit model was chosen because it slightly outperforms its probit counterpart in terms of the in-

Apparently, lower secondary school GPAs contribute to the prediction of unexcused absences in vocational school and are actually the only valid predictors in the reduced model. In the full model, the coefficient of the type of school is also significant. Figure 5 illustrates the predictive power of these two variables. An applicant who belongs to the reference group of the regression applying with a GPA of 4.0 has a considerably higher probability of having unexcused vocational school absences than an applicant from an intermediate-level school with a GPA of 5.0, *ceteris paribus*. However, overall Multicheck scores have no predictive power concerning the unexcused absences of apprentices.

**Figure 5. Probability of unexcused vocational school absences according to lower secondary school GPAs and types of schools**



It is not negligible that Multicheck fails to predict unexcused absences in vocational school: disciplinary reasons are the main determinants of premature apprenticeship terminations at Migros (according to the trainers of apprentices) - which is, arguably, the most evident sign of failure in vocational training. Can the test nonetheless

formation criteria. One might fear that the logit regressions suffer from clustering. This fear is reasonable because the probability of having unexcused absences significantly differs across vocational schools. Therefore, several logit and probit regressions-school-fixed-effects logit, and random-effects probit-were conducted to control the results. The main coefficients are nearly identical across models. Furthermore, the standard deviations of the main coefficients are even smaller than those given in Table 4 if we use a cluster-robust estimator.

predict a premature apprenticeship termination? This result would be desirable, as the test assesses the «aptitude» of an adolescent for a certain apprenticeship.

Unfortunately, it was not possible to collect the data for all of the apprentices who began their apprenticeships in mid-2007 or 2008 but later dropped out. As a result, the data set contains only fourteen drop-outs in a sample of 250 apprentices (see Appendix B for a discussion of this issue). Nevertheless, the binary variable indicating a drop-out was regressed on the overall Multicheck scores, the lower secondary school GPAs, and some control variables.<sup>13</sup>

The results of the probit regression are shown in Table 3. These results are not fully reliable because they depend relatively strongly on individual drop-outs and because they are based on an incomplete sample. However, two results were robust in all specifications, subsamples, and models tested: the overall Multicheck scores are not able to predict the probability of premature apprenticeship terminations, but the lower secondary school GPAs are able to predict these terminations (at least on a significance level of ten percent). The latter finding is notable because only four of the fourteen drop-outs were explicitly caused by problems in meeting the academic requirements of the vocational school. All other premature apprenticeship terminations were caused by disciplinary problems, insufficient practical performance, or incorrect choices of occupations.

Table 3: Determinants of premature apprenticeship terminations  
(Probit regression)

Dependent variable	Premature apprenticeship termination
Overall score MC	0.021 (0.0173)
GPA	-0.632+ (0.332)
Female	0.325 (0.301)
German	0.191 (0.311)
Urban	0.315 (0.325)
Intermediate-level school	-1.148** (0.443)
Constant	0.281
Observations	250
McFadden's Pseudo- R <sup>2</sup>	0.116

<sup>13</sup> Some explanatory variables of the regressions above were excluded, as they were insignificant and did not influence the results of the main coefficients.

## 7. Concluding remarks

Lower secondary school GPAs and the types of schools the applicants attended are reliable sources of information for firms in search of apprentices. The signalling value of school grades in this study is remarkable, as earlier work has indicated that school grades tend to be systematically biased and noisy, and grades are an imperfect indicator of signal scholastic or cognitive abilities. Why could school grades nevertheless be valid predictors of vocational training success? First, the result can be explained with the reduction in the noise and bias in grades when they are averaged across individual tests, subjects, and teachers (cf. Grant, 2007). Secondly, studies have shown that grades contain social components that depend on the character (above all, frankness, friendliness and sense of duty), interest, and motivation of the pupils (Baron-Boldt, 1989; Imdorf, 2007; Lekholm & Cliffordson, 2009; Miller, 1998). The positive link between the school grades and non-cognitive skills of pupils may also explain the relatively strong predictive power of school grades in this study.

However, the Multicheck retail sale test cannot meet the expectations of firms. The test is not a valid predictor of the training success of apprentices at Migros. A substantial part of the small but positive correlation between overall Multicheck scores and vocational school grades seems to stem from the tendency of higher test results for apprentices who attended higher-level schools-which is information that the firm can infer from the application dossier. In addition, the test does not contribute to the prediction of inappropriate social behaviour (the probability of unexcused absences in vocational school) or the likelihood of premature apprenticeship termination. If one is willing to believe that these results do not strongly suffer from a sample selection bias, then we can conclude that the test does not significantly improve Migros' selection of apprentices.

Are the findings of this investigation transferable to other circumstances, apprentices, and firms; that is, are they externally valid? In this context, it is important to note that the results contain a substantial amount of variation in individual training and learning situations and are hence valid in very different settings. The apprentices come from twelve different cantons and dozens of different lower secondary schools with different types of schools, and they are educated in 23 different vocational schools, in three different cooperatives, and in nearly hundred individual branch offices. Nevertheless, we cannot be certain that the results are generalisable to other firms, particularly because studies have shown that vocational success is considerably firm-specific (e.g. Stalder & Schmid, 2006; Stamm et al., 2010) and that firms differ with respect to their recruitment of apprentices (Imdorf, 2009; Moser, 2004). Thus, the results of this study might be specific to the context of the Migros cooperatives examined.

However, even if the results are only valid for the Migros cooperatives, they are still important. First, Migros represents more than one-third of the Swiss retail business. Secondly, the weak predictive validity (as, e.g., illustrated in Figure 2) and the

conceptual problems mentioned strongly suggest that the test fails to measure the «aptitude» of test takers as retail sales specialists. This conclusion can be drawn irrespective of possible sample selection or firm effects: as the test aims to indicate whether an adolescent is able or unable to complete an apprenticeship as a retail sales specialist, the test should be diagnostically correct regardless of the specific context of the firm or of unobservable qualities - or it should measure them. In conclusion, the test seems to represent more of an assessment of the knowledge taught at school than a proper «aptitude» test for apprentices.

## References

- Amos, J., Amsler, F., Martin-Jahncke, M. & Michel, B. (2009). Evaluation der Resultate von Lehrabschlussprüfungen der beruflichen Grundausbildungen. Büro für Kommunikation, Basel
- Arrow, K. J., (1973). Higher education as a filter. *Journal of Public Economics*, 2(3), 193-216.
- Baron-Boldt, J. (1989). Die Validität von Schulabschlussnoten für die Prognose von Ausbildungs- und Studierenerfolg: eine Metaanalyse nach dem Prinzip der Validitätsgeneralisierung, Frankfurt a. M.: Lang
- Bauer, P. & Sheldon, G. (2008). *Ethnic Discrimination in Education: The Swiss Case*, Basel: University, Department of Economics, FAI
- Bertschy, K., Böni, E. & Meyer, T. (2008). Young people in transition from education to labour market. Results of the Swiss youth panel survey TREE, update 2007. Basel: TREE
- Bettinger, E. P., Evans, B. J. & Pope, D. G. (2011). Improving college performance and retention the easy way: unpacking the ACT exam, NBER Working Paper No. 17119
- BPS (2005). PISA 2003: Kompetenzen für die Zukunft. Zweiter nationaler Bericht. Neuenburg: BFS
- Bishop, J. H. (1994). Signaling the competencies of high school students to employers. CAHRS Working Paper 94-18
- Costrell, R. M. (1994). A simple model of educational standards. *The American Economic Review*, 84(4), 956-971
- De Paola, M. & Scoppa, V. (2010). A signalling model of school grades under different evaluation systems. *Journal of Economics*, 101(1), 1-14
- Fielding, A., Yang, M. & Goldstein, H. (2003). Multilevel ordinal models for examination grades. *Statistical Modelling*, 3(2), 127-153
- Grant, D., (2007). Grades as information. *Economics of Education Review*, 26(2), 201-214
- Häberlin, U., Imdorf, C. & Kronig, W. (2004). Von der Schule in die Berufsschule. Untersuchungen zur Benachteiligung von ausländischen und von weiblichen Jugendlichen bei der Lehrstellensuche. Bern: Haupt
- Heckman, J. & Vytlačil, E. (2001). Identifying the role of cognitive ability in explaining the level of and change in the return of schooling. *The Review of Economics and Statistics*, 83(1), 1-12
- Hunsley, J. & Meyer, G. J. (2003). The Incremental Validity of Psychological Testing and Assessment: Conceptual, Methodological, and Statistical Issues. *Psychological Assessment*, 15(4), 446-455
- Imdorf, C. (2005). Schulqualifikation und Berufsfindung. Wie Geschlecht und nationale Herkunft den Übergang in die Berufsbildung strukturieren, Wiesbaden: Verlag für Sozialwissenschaften
- Imdorf, C. (2007). Die relative Bedeutsamkeit von Schulqualifikationen bei der Lehrstellenvergabe in kleineren Betrieben. In: T. Eckert, (Hrsg.), *Übergänge im Bildungswesen*. Münster: Waxmann
- Imdorf, C. (2009). Die betriebliche Verwertung von Schulzeugnissen bei der Ausbildungsstellenvergabe. *Empirische Pädagogik*, 23(4), 392-409
- Kronig, W., (2007). Die systematische Zufälligkeit des Bildungserfolgs. Bern: Haupt
- Lekholm, A. K. & Cliffordson, C. (2009). Effects of Student characteristics on grades in compulsory school. *Educational Research and Evaluation: An International Journal on Theory and Practice*, 15(1), 1-23
- Lindahl, E. (2007). Comparing teachers' assessments and national test results - evidence from Sweden. IFAU - Institute for Labour Market Policy Evaluation

- Lounsbury, J. W., Steel, R. P., Loveland, J. M. & Gibson, L. W. (2004). An investigation of personality traits in relation to adolescent school absenteeism. *Journal of Youth and Adolescence*, 33(5), 457-466
- Miller, S. R. (1998). Shortcut: high school grades as a signal of human capital. *Educational Evaluation and Policy Analysis*, 20(4), 299-311
- Moser, C., Stalder, B. E., & Schmid, E. (2008). *Lehrvertragsauflösungen: Die Situation von ausländischen und Schweizer Lernenden. Ergebnisse aus dem Projekt LEVA*, Bern: Erziehungsdirektion des Kantons Bern
- Moser, U. (2004). *Jugendliche zwischen Schule und Berufsbildung. Eine Evaluation bei Schweizer Grossunternehmen unter Berücksichtigung des internationalen Schulleistungsvergleichs PISA*, Bern: hep
- Mühlemann, S., Wolter, S., Fuhrer, M. & Wüest, A. (2007). *Lehrlingsausbildung - ökonomisch betrachtet. Ergebnisse der zweiten Kosten-Nutzen-Studie*, Zürich: Rüegger
- Mühlemann, S., Wolter, S. C. & Wüest, A. (2009). Apprenticeship training and the business cycle. *Empirical Research in Vocational Education and Training*, 1(2), 173-186
- Multicheck, (2010). *Multicheck - Wissenschaftliche Erhebung*, [http://multicheck.potentials.ch/fileadmin/user\\_upload/Wissenschaftliche\\_Dokumente/Broschuere\\_Wissenschaft.pdf](http://multicheck.potentials.ch/fileadmin/user_upload/Wissenschaftliche_Dokumente/Broschuere_Wissenschaft.pdf). Konolfingen: Multicheck
- Murnane, R. J., Willett, J. B., Braatz, M. J. & Duhaldeborde, Y. (2001). Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY. *Economics of Education Review*, 20(4), 311-320
- Schmidt, F. L. & Hunter, J. E. (1998). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin*, 124(2), 262-274
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355-374
- Stalder, B. E., Meyer, T. & Hupka-Brunner, S. (2008). Leistungsschwach - Bildungsarm? Ergebnisse der TREE-Studie zu den PISA-Kompetenzen als Prädiktoren für Bildungschancen in der Sekundarstufe II. *Die Deutsche Schule*, 100(4), 436-448
- Stalder, B. E. & Schmid, E. (2006). *Lehrvertragsauflösungen, ihre Ursachen und Konsequenzen. Ergebnisse aus dem Projekt LEVA*. Bern: Erziehungsdirektion des Kantons Bern
- Stamm, M., Niederhauser, M. & Kost, J. (2010). The top performers in vocational training. *Empirical Research in Vocational Education and Training*, 2, 65-81
- Stiglitz, J. E. (1975). The theory of «screening», education, and the distribution of income. *The American Economic Review*, 65(3), 283-300
- Sund, K. (2009). Estimating peer effects in Swedish high school using school, teacher, and student fixed effects. *Economics of Education Review*, 28(3), 329-336
- Weiss, A. (1995). Human capital vs. signalling explanations of wages. *Journal of Economic Perspectives*, 9(4), 133-154
- Widmer, M. (2006). *Reliabilität und Validität eines Berufseignungstests bei kaufmännischen Lehrlingen*, Bern: Universität Bern

Appendix

A. Multicheck evaluation sheet



Info@multicheck.ch  
www.multicheck.ch  
031 791 01 16

Multicheck Detailhandel und Service

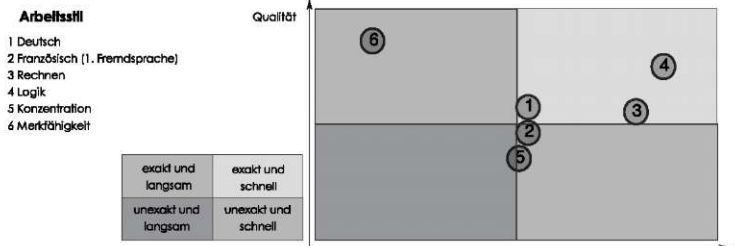
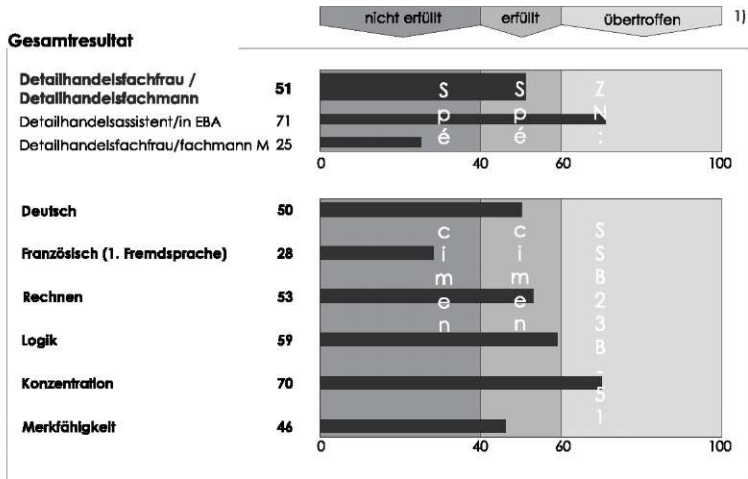
Die Ergebnisse sind stricthrechtlich geschützt.

2010/11

Durchführung: 30.04.2010  
Name: Spécimen  
Strasse: Beispielstrasse 2  
Geburtsdag: 20 April 1995

Vorname: Spécimen  
PLZ: 3000  
Ort: Bern

Ergebnisse  
ZN: SSB23B-51



Version 01

1) Siehe beiliegende Broschüre oder auf dem Internet: www.multicheck.ch

Ein gutes Ergebnis garantiert nicht für eine Lehrstelle; Zeugnisse, Gespräche, Angebot und Nachfrage ... sind ebenso wichtig!

**B. Definitions of variables and coding***Variables from the CVs*

CBGT: dummy variable indicating whether an apprentice spent a transitional year in a course to bridge gaps in training.<sup>14</sup> Here, a CBGT is any employment experience or course passed during a transitional year that is likely to have increased the scholastic abilities that are beneficial for vocational school.

Gender: <16 yrs, >17 2/3 yrs: dummy variables indicating whether an apprentice belongs to the highest (older than 17.67) or lowest (younger than 16) quintile of the sample age distribution of apprentices at the beginning of the apprenticeship. The use of dummy variables rather than a continuous variable is motivated by the presumption of nonlinear age effects.

German: dummy variable indicating whether or not the native language of an apprentice is German  
Skill of parents: dummy variable that acts as a proxy for the social background of an apprentice. This variable is based on information regarding the professions of parents available in most CVs. These professions were coded according to the Austrian version of the International Standard Classification of Occupations (Ö-ISCO): The classification groups all professions hierarchically into ten different categories, which in turn correspond to four skill levels. The variable is based on these skill levels. The dummy takes the value one if a profession requires a skill level of three or four. A higher skill level of either parent is decisive for the assignment. The variable is not part of the standard set of control variables used in the regressions because it could be coded for only 247 apprentices and had no major effect on the results of the regressions.

*Variables from lower secondary school reports*

GPA: Average grade derived from the newest and second-newest lower secondary school reports. Only grades from mathematics, German, French, English, history and geography (taken together), and natural sciences are considered. Grades range from 6.0 (excellent) to 1.0 in Switzerland, commonly in steps of 0.5. The lowest sufficient grade is 4.0. I always consider grades from the two most recent lower secondary school reports regardless of whether these reports were derived from the eighth, ninth, or tenth school year. The comparison of grades across school years is justified because firms hire apprentices based on their most recent school records.

Type of school (basic-, intermediate-, or advanced-level school): The types of schools were classified into three different performance levels, which were similar to those used in the publications of the Swiss Federal Statistical Office (e.g., BFS, 2005). However, this classification is a simplification. Some cantons in Switzerland place pupils at the end of the sixth (or fourth) school year into two or four different types of schools according to their performance levels. In addition, school systems also vary in the amount of permeability of the school levels.

Urban: dummy variable indicating whether the lower secondary school the apprentice attended was situated in an urban or rural area. The variable was coded using the spatial typology of the Swiss Federal Statistical Office: This typology groups regions into four spatial categories: cities (1), agglomerations (2), and rural regions (categories 3 and 4). Consequently, any school situated in an area with an index that is less than or equal to two is treated as 'urban'.

*Variables from the Multicheck evaluation sheet*

Overall score MC: overall Multicheck score ranging from 1 to 100

Test scores of subsections of Multicheck ranging from 1 to 100: German, French, English, numeracy, logic, concentration, retentiveness

*Variables from vocational school reports*

Dummy variables for the Migros cooperative for which an apprentice works (Aare, Lucerne, and Zurich)  
Unexcused absences: dummy variable that is one if an apprentice had unexcused absences from vocational school during the first two years of the apprenticeship. This variable is binary rather than continuous because some vocational schools do not provide information regarding absences in the school records. The apprentice trainers provided the missing information for the apprentices from Migros Zurich, but the information was provided only in binary form. Nevertheless, the sample size reduces to 290 because the information regarding absences was unavailable for some apprentices from Migros Aare and Lucerne.

Premature apprenticeship termination: dummy variable indicating whether an apprenticeship contract

<sup>14</sup> These courses are offered as transitional solutions for young people who, after completing the lower secondary level, do not immediately begin basic vocational education and training or do not continue their education at an upper secondary school.



was terminated prematurely. Unfortunately, the application material from several drop-outs had already been returned to the adolescents; thus, it was impossible to collect the data from all drop-outs. Furthermore, Migros Lucerne did not provide information pertaining to their drop-outs because of privacy concerns. All apprentices from this cooperative have been excluded from the analysis. Vocational school GPA 1 and vocational school GPA 3: mean vocational school grades in the first and third semesters, respectively, based on grades in all subjects taught at vocational school: economy, a foreign language (either French or English), general aspects of the branch (Branchenkunde), German, society, and retail sales skills (Detailhandelskenntnisse). The grade from the subject retail sales skills is weighted twice in the calculation of the GPA because the subject counts as double in the final exam of the apprenticeship.

### C. Logit regressions of unexcused vocational school absences

**Table A5: Determinants of unexcused vocational school absences  
(Logit regressions)**

Dependent variable	Unexcused vocational school absences	
	(1)	(2)
Overall score MC	0.0223 (.0197)	0.0114 (0.0218)
GPA	-0.9313* (0.3941)	-1.041* (0.460)
>17 2/3 yrs	0.233 (0.446)	0.304 (0.478)
<16 yrs	-0.0474 (0.6874)	0.218 (0.733)
Female	-0.4748 (0.344)	-0.520 (0.370)
German	0.0023 (0.389)	0.0229 (0.409)
Urban	0.455 (0.389)	0.545 (0.412)
CBGT	0.0425 (0.386)	-0.145 (0.412)
Intermediate-level school	-0.604 (0.389)	-0.995* (0.477)
Advanced-level school	0.6469 (0.9626)	-0.160 (1.102)
Constant	1.4869	3.877*
Cantonal dummies	No	Yes
Dummies for cooperatives	No	Yes
Observations	290	290
McFadden's Pseudo-R <sup>2</sup>	0.0768	0.166

\*\* p < 0.01, \* p < 0.05, + p < 0.1

Notes: Standard errors in parentheses.