Can we identify the students who have success in macroeconomics depending on exam format by comparing multiple-choice test and constructed response test?

Leiv Opstad

Norwegian University of Science and Technology (NTNU),
7491 Trondheim, Norway
Email: leiv.opstad@ntnu.no

Abstract: The selection of assessment format in business schools and higher education has been a topic of consideration for many years. Currently, in Norway, there is a debate about replacing constructed response (CR) questions with multiple-choice (MC) questions. MC tests are popular and have been substituted for traditional exams in many fields. In the context of costs, there is obviously a gain to using MC tests. By exploring data from a macroeconomics course that applied both essay-based and multiple-choice-based tests simultaneously, it was possible to compare the two exam formats and to identify students who performed differently. By using a linear regression model, the findings revealed that there were substantial differences in the influence of the independent variables for the two methods. For CR, the result confirms prior research, while there was almost no significant connection between the chosen independent variables and the dependent variable MC. Therefore, students' background, skills and personal characteristics matter.

1   Introduction

In recent years there has been pressure within higher education and especially within business schools to replace traditional essay or constructed response (CR) tests with multiple-choice (MC) tests, especially in introductory courses (Simkin and Kuechler,

2005). By selecting MC, fewer resources are required, the tests are easy to implement, and students get quick feedback (Douglas et al., 2012). Computers can calculate the score, and there is no bias in the grading practice. An online MC-based exam gives flexibility, and students' progress can easily be tracked over time by having more than one test during the course. However, there are some pedagogical limitations when using MC. The two test methods (CR and MC) might evaluate different kinds of knowledge, and it is easier to measure in-depth learning with CR (Simkin and Kuechler, 2005). Moreover, the use of MC tests makes it hard to appraise the higher understanding of learning by applying the definition of Bloom's taxonomy (Anderson et al., 2001; Buckles and Siegfried, 2006).

The purpose of this study is to identify those students who achieve different performance depending on the selected exam form. This will be done by analysing different factors (students' personality traits, gender and skills) explaining students' assessments by comparing the results from MC and CR for business students taking both tests simultaneously on the same course at Norwegian University of Science and Technology (NTNU). The number of students attending such compulsory courses is over 200. In the exams, CR questions are based on written assignments containing essays and analyses. There are mostly open-ended questions that evaluate some difficult reasoning and some higher level of understanding. We wanted to see if there were different impacts among the independent variables. Prior research indicates that gender, mathematical skills, personal characteristics and students' effort all have influence on success in the study of macroeconomics (Opstad, 2018; Opstad and Fallan, 2010). Do the same students perform well regardless of the exam format? If not, what characterises those students who achieve different results? Due to coordination with other colleges, the number of students in compulsory courses at NTNU Business School has doubled (to over 500) in the last year. Therefore, NTNU is considering applying more MC tests. It is not easy to give a simple answer as to what impact this will have on students' success in business courses. By analysing factors correlated with the exam results by applying both MC and CR questions, we can gain a little more insight and make some contribution to this discussion. This study shows that there is some variation among the factors that are associated with the performance depending on the chosen type of test. More research needs to be carried out on this topic. This paper does not discuss in detail why there might be differences between CR and MC, or whether substitution of CR with MC will have an impact on the way undergraduates study and acquire knowledge in economics. The intention of this article is not to compare the two test formats by measuring the validity and reliability.

2   Theory and previous literature

2.1   Gender and performance in economics courses

Numerous studies indicate that men get higher scores in college economics (microeconomics and macroeconomics) than do women (Anderson et al., 1994; Watts, 1987). The finding of a gender gap in economics is not universal. Some researchers have not reported any gender difference (Parker, 2006; Swope and Schmitt, 2006). However, the tendency is for men to perform better than women in collegiate economics courses. In a meta-analysis including 68 studies published since 1980, Johnson et al. (2014) reported

that men outperformed women in 68% (31% was statistically significant) of the cases reviewed. In 9% of the studies, men performed significantly worse than women. There might be many reasons for the gender gap in achievement in macroeconomics and other economics courses, for instance gender differences in personal characteristics. By including those factors in the analyses, Opstad and Fallan (2010) reported a lower gender impact in macroeconomics performance. Furthermore, the gender gap is closing as time goes on (Johnson et al., 2014).

2.2 *Mathematical skills and performance in economics courses*

Several researchers testify that mathematics ability has an impact on students' success in the principles of economics courses (Arnold and Straten, 2012; Brown-Robertson et al., 2015; Mallik and Lodewijks, 2010; Opstad, 2018). According to Ballard and Johnson (2004), quantitative skills are probably the key factor for success in economics courses. The students' mathematical background is vital since quantitative skill is an important tool for analysis, especially in microeconomics but also in macroeconomics. There is a strong link not only between mathematical skills in performance on economics courses, but also in all quantitative business courses (Opstad, 2018). The conclusion is that a mathematical skills background is a good predictor of academic performance in studies of economics and business (Arnold and Rawaan, 2014).

2.3 *Students' effort and performance in economics*

Study effort is also a crucial determinant of academic success. There is a correlation between students' effort and performance in economics courses (Bonesrønning and Opstad, 2012; Stinebrickner and Stinebrickner, 2008). Extra study time leads to better grades. However, other studies have not found such an effect (Guillaume and Khachikian, 2011). Sansani and Rahamin (2019) suggest there are some unobserved factors to explain this. Such issues can be academic skills, preferences, and personal characteristics.

2.4 *Students' attendance and performance in economics*

Some studies find a positive association between class attendance and performance in economics courses (Cohn and Johnson, 2006; Horn and Jansen, 2009; Opstad and Fallan, 2010). However, the overall impact does not seem to be substantial. For the average student, absence from lectures has only a minor influence on the learning standard. In general, the literature shows a positive but mainly weak relationship between attendance and achievements (Büchele, 2020). Controlling for other key variables (like the professor's teaching style, academic skills, motivation, interest, and study effort including homework), some authors find only a weak or no link between attendance and success in the study fields (Rodgers, 2001).

2.5 *GPA and performance in economics*

A number of studies show a positive correlation between grade point average (GPA) during upper secondary school and performance in economics at college and university (Cannonier and Smith, 2019; Kassis and Boldt, 2020; Opstad and Fallan, 2010). With

high GPA scores from upper secondary school, the students seem to be able to solve problems in many business and economics courses (Brookshire and Palocsay, 2005). The admission to business schools is based on students' GPA marks from upper secondary schools. Therefore, it makes sense that there is a link between GPA scores and success in business courses. However, Opstad (2018) suggests GPA may not be a good indicator for predicting success in economics courses. Other factors like mathematical skills seem to play a more important role in explaining results in macroeconomics.

2.6 *Personal characteristics and academic performance*

One of the most commonly used methods for measuring personal characteristics is the big five model (Costa and McCrae, 1992). It consists of five factors: agreeableness, conscientiousness, neuroticism, extraversion and openness. Agreeableness is related to persons who behave harmoniously, participate and are pleasant. Conscientiousness is linked to working hard, being orderly and well organised. Neuroticism is the opposite of emotional stability. Neurotic people often experience stress, depression and anxiety. Extraversion is defined as being expressive, social and outgoing. An extraverted person likes to be together with others. Openness describes people who are open to experience, changes and new ways of learning.

Table 1   The big five personality traits

Trait	Definition. The person tends to:
Openness to experience	Be open to new aesthetic, cultural or intellectual experiences
Conscientiousness	Be organised, responsible and hardworking
Extraversion	Be interested and focussed on the outer world of people and things rather than the inner world of subjective experience
Agreeableness	Act in a cooperative, unselfish manner
Neuroticism (inverse of emotional stability)	Have a chronic level of emotional instability and a proneness to psychological distress

Several researchers have investigated the relationship between personality traits and academic success. Important contributors to this discussion are O'Connor and Paunonen (2007) and Trapman et al. (2007). The factor that is the best predictor of success at universities is conscientiousness. There is a substantial link between conscientiousness and academic achievement, with small differences for different majors, cultures and countries. Some argue that conscientiousness is the only predictor trait for academic achievement (Buju, 2013; Ibrahim et al., 2014). The explanation is that students who are well organised, efficient, systematic and goal-oriented will perform well. Most of the research does not find neuroticism and agreeableness to be important determinants for academic success. Business students score higher in emotional stability than other students, but this does not seem to have an impact on academic performance (Lounsbury et al., 2009). The relationship between extraversion and performance is not clear, since prior findings are mixed. O'Connor and Paunonen (2007) suggest a weak negative correlation between extraversion and achievement. In some publications the authors report a positive association between extraversion and performance (Chowdhury, 2006; Fallan and Opstad, 2014; Rothstein et al., 1994). It is also noted by some authors that students enrolled in business and economic-oriented fields get high scores in extraversion

and conscientiousness (Matthews and Oddy, 1993). Contrary to what one would expect, openness is not a good predictor for academic success. Many researchers fail to show a significant positive link (O'Connor and Paunonen, 2007). Since business students are in general more conservative than other students and less open to experience, it is not obvious how the connection can be established (Lounsbury et al., 2009). De Fruyt and Mervielde (1996) and Busato et al. (2000) report a negative correlation between openness and performance.

2.7 The relation between exam form (CR or MC) and performance

It is challenging to make an MC-based test that involves higher-order thinking skills (Aiken, 1982). However, exam design can be more easily accomplished by presenting CR-based questions (Hickson et al., 2012). In order to promote reflective judgement, analysis and evaluation, as part of an exam this can be easily accomplished by using essays (Dwyer et al., 2014). Students can use their own words, analyse the topics, and express originality. Therefore, MC questions and essays can substitute for each other at the undergraduate level since both methods are substantially effective at capturing students' achievement at a less-advanced level (Becker and Johnson, 1999). If the two types of exams measure different kinds of knowledge, the connection between students' scores in MC questions and essay questions is weak. Some researchers confirm this. With data from a programming class examining different levels of knowledge, Kuechler and Simkin (2010) do not find a close relationship between these two testing methods. Bridgeman (1992) reports that MC formats were nearly overlapping for some items, but not for others. MC questions measure only some parts of the students' knowledge. Bridgeman concluded that an open-ended formatted exam is superior to MC questions. MC-based questions cannot achieve the same advanced level of knowledge as CR. However, this depends to some degree on how the MC questions are designed (Chan and Kennedy, 2002).

There are also some similarities between using MC or essays as an exam format. Walstad and Becker (1994) report a correlation coefficient of 0.65 in macroeconomics and 0.69 in microeconomics. The scores tend to be higher in quantitative than in non-quantitative courses (Hickson et al., 2012). Since the correlation coefficient between two CR-based tests within the same course is almost identical to that recorded between MC and CR, Hickson et al. (2012) argue that the outcome of changing from CR format to MC format will be minimal. The authors claim that independently of the test method, there is no advanced analysis in introductory courses at the undergraduate level. The highest levels of academic skill (Bloom's taxonomy) are normally suitable for master's degree and PhD-level students. If one wants to measure advanced dimensions of learning and knowledge, there will be a considerable gap between those two methods.

3 Hypotheses

Based on theory and previous research, we postulate the following hypotheses by comparing MC and CR:

Hypothesis 1 (H1)	There are differences between the students performing well depending on exam format in macroeconomics.
Hypothesis 2 (H2)	The correlation between gender and performance in macroeconomics differs depending on exam format.
Hypothesis 3 (H3)	The link between mathematical skills and scores in macroeconomics differs depending on exam format.
Hypothesis 4 (H4)	The relationship between students' effort and their results in macroeconomics depending on exam format.
Hypothesis 5 (H5)	The relationship between students' attendance and their results in macroeconomics differs depending on exam format.
Hypothesis 6 (H6)	The relationship between students' GPA and their performance in macroeconomics differs depending on exam format.
Hypothesis 7 (H7)	Personality traits have an influence on performance in macroeconomics depending on exam format.

Some previous research has shown a relationship between performance in MC and CR. Hypothesis 1 tests if the influence varies depending on exam format.

Based on the international literature, it has been documented that there is a positive link between performance in macroeconomics and the following factors: gender, mathematical background, study effort, and GPA. As a result of these reports, we postulate Hypotheses 2, 3, 4 and 6. Furthermore, previous studies have verified that personal characteristics have an influence on performance (H7).

Prior research has also showed a positive correlation between attendance and performance, but the relationship is rather weak. There might be a different impact between MC and CR.

If there are different scores depending on the exam format, this is an indication that different students have success depending on the exam format.

4   Sample and research methodology

4.1   Sample

The sample comprised 229 students during the years 2018 and 2019 (Table 2a) at NTNU. The questionnaire was distributed among undergraduates at the NTNU Business School who were present in the compulsory macroeconomics class; since 30% of the students were normally absent, the data were non-randomly selected. Therefore, the data were not representative of the final sample. However, earlier research shows small differences between the sample and the entire population (Bonesrønning and Opstad, 2015). We presumed this was the case in this study. Like Park and Kerr (1990), we measured students' effort as out-of-class study hours. This is a proxy, since it does not say anything about how effective this activity is. The students reported an average self-study of four hours per week. In addition, the participants reported on the number of lectures they had attended. Altogether, there were 15 lectures in this course. The average presence at lectures was quite high (12.5).

Table 2a Descriptive statistics from the survey

	N	Minimum	Maximum	Mean	Std. deviation	Scale reliability
Cronbach's alpha						
Study effort (hours a week, lectures excluded)	226	0	15	3.98	2.40	
Attendance	226	2	15	12.52	2.07	
Expected performance[1]	221	1	5	3.55	0.74	
Personality traits[2]						
Extraversion	229	1.50	5.00	3.65	0.77	0.805
Agreeableness	229	2.00	5.00	3.98	0.54	0.636
Conscientiousness	229	2.00	5.00	3.78	0.62	0.632
Emotional stability	229	1.50	5.00	3.31	0.73	0.694
Openness	229	1.25	5.00	3.23	0.77	0.674

Notes: [1]F:0, E:1, D:2, C:3, B:4, A:5 and [2]Likert scale 1: strongly disagree, 5: strongly agree.

Table 2b Descriptive statistics (percent) from administrative data

N	Variable	Percent
196	Females	54.0
229	N-mathematics	35.0
229	S-mathematics	42.0
229	P-mathematics	23.0

Table 2c Descriptive statistics about performance from administrative data

	N	Minimum	Maximum	Mean	Std. deviation
GPA (from upper secondary school)[1]	155	46.90	66.70	51.34	2.99
Letter grade (final exam)[2]	229	1	5	3.24	1.20
ET score (0–75 points)	229	4.00	70.00	42.12	13.05
MC score (0–25 points)	229	1.00	23.00	13.19	4.15

Notes: [1]grades from outstanding to fail: 6, 5, 4, 3 and [2]multiplied by ten plus additional points F:0, E:1, D:2, C:3, B:4, A:5.

Output expectation at the final exam is a proxy for self-efficacy (Fallan and Opstad, 2016). Self-efficacy expresses how individuals evaluate their capability to perform courses of action needed to achieve designated types of performance (Bandura, 2010). The mean values of personality traits are quite high. The reliability coefficients (Cronbach's alpha) are at accepted levels.

The information from the survey was connected to administrative information about the students' background and actual performance in the final exam in macroeconomics (Tables 2b and 2c).

Mathematics is compulsory in the first two years at upper secondary school. The students can choose between practical mathematics (P-maths), mathematics for business and social sciences (S-maths) or mathematics for natural sciences (N-maths). They also have opportunities for further studies in S- and N-maths in the third year. All three pathways in mathematics from upper secondary school qualify for admission to business studies courses.

Note that 35% of the undergraduate business students chose N-maths at upper secondary school. Furthermore, 42% selected the S-maths pathway, which is adapted to business studies. The remaining students took P-maths (23%). There are more female students (54%) than males (46%). Since there is great competition to gain entry to this business school, students from upper secondary school have consistently good grades in all subjects. Regression analysis related to the study of Opstad et al. (2017) did not show any significant correlation between the grades in mathematics at secondary school and performance in business courses.

4.2   The assignment format in macroeconomics and outcomes

One of the purposes of the exam is to measure to what extent the students have achieved the learning goals in the subject.

Learning goals in this macroeconomics course are to:

a   get knowledge about the subject

b   be able to utilise quantitative and analytical methods

c   apply and use theory, for instance the impact of changes in variables like the exchange rate

d   demonstrate active reflection in the subject.

The intention is to measure the same dimensions in the final exam regardless of the test method (practical knowledge, calculation, analysis and application of theory). The students have an individual written four-hour exam which decides the letter grade. The exam is a combination of 5 to 10 CR questions and 32 MC questions. The total score is 100, with 75 from CR and 25 from MC. One needs less time to answer MC questions than CR questions. Therefore, a pure MC-format test tends to take less time than a pure CR-format test. Examinees must be aware of this when evaluating and comparing the two test methods. For each MC question, there are five alternatives. By just guessing, a student achieves an average of 6.4 in score (i.e., 20% correct of 32 possible). There is no negative mark for incorrect answers. Answers based on only guessing do not help, since the participants only get benefit if there are more than seven correct answers. The letter grade depends on the total score from those two exam formats. The intention with both tests is to measure all four learning goals (a–d). It is challenging to measure students' ability for active reflection (point d) by using MC questions. The unique circumstance with this study is that one can compare the results of two exam formats for the same subject simultaneously for the same group. Furthermore, the administrative data (regarding demographics) can be linked to information from the survey.

Table 3   Correlation (Pearson) matrix

Item	MC-score	CR-score	Expected result
MC-score	1.0	0.516	0.334
CR-score	0.516	1.0	0.454

The result confirms prior research. There is a strong correlation between MC-format exams and CR-format exams. The connection between the two designs and the students' expected performance (proxy for self-efficacy) is weaker and the correlation between expected letter grade and MC score is substantially lower than the link between expected letter grade and CR score. This might be an indication that the two methods do not measure the same type of knowledge. Students can be aware of the difference between the two tests. Since CR questions are emphasised more than MC questions in determining the final grade, students give CR issues a higher priority in their preparation for the exam. Hence, this can also explain why the correlation coefficients will be different for the two test methods. If so, this verifies that the two methods measure diverse dimensions.

To make it easier to compare the outcomes of the two test formats, we have applied an equal scale from 0 to 100 (dividing MC scores by 0.15 and CR scores by 0.75, it gives both a scale from 0 to 100). Figures 1, 2 and 3 present the results. The mean score is highest for the CR-format exam (the mean difference is 7.5), but the variation is almost the same for both test methods. This is in line with the finding of Hickson et al. (2012). The two assessment designs produce the same grade of variations. Even though there is a strong correlation between the two methods, some students perform well in one of the tests but not in the other one. Note that the standard deviation is also substantial for the difference between the two exam patterns (see Figure 3). This confirms to some degree Hypothesis 1 (H1). Some students perform much better in an MC test than a CR test, and vice versa, while others achieve about the same outcome independent of test method.
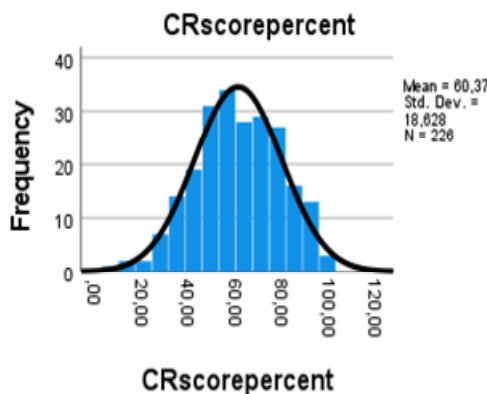
Figure 1   Distribution CR-format test (scale 0–100) (see online version for colours)

Figure 2   Distribution MC-format test (scale 0–100) (see online version for colours)
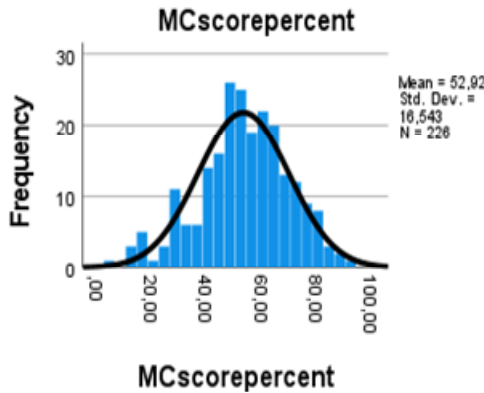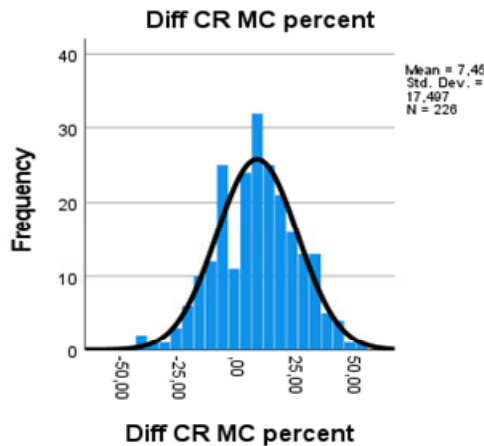

Figure 3   Distribution difference between CR- and MC-format test (scale 0–100)
(see online version for colours)

The skewness and kurtosis measurements show that all three figures are distributed in relatively symmetrical and normal distribution (Table 4).

Table 4   Skewness and Kurtosis

Test-method	Skewness	Kurtosis
MC	–0.328	0.028
CR	–0.166	0.368
CR-MC	–0.194	–0.073

4.3   The regression model

For exploring more the variation in performance among the students depending on the exam format, we will introduce a regression model. If the two methods have the same impacts, the link between the explanatory variables and outcomes should be the same independent of exam format. Deploying a linear regression model makes it possible to

simultaneously determine the different factors that are associated with the score in macroeconomics using the two test methods.

The model's dependent variable is performance in macroeconomics.

$$Yi = a_0 + a_1 X_1 + a_2 X_2 + a_3 X_3 + a_4 X_4 + a_5 X_5$$
$$+ a_6 X_6 + a_7 X_7 + a_8 X_8 + a_9 X_9 + a_{10} X_{10} + \varepsilon$$

where

- Yi is the student's score in the final exam in macroeconomics i:
  1   is the MC test (score 0 to 25)
  2   is the CR test (score 0 to 75).

- a_0 is a constant

- X_1 is gender (0:F, 1:M)

- X_2 is a dummy variable for N-mathematics (0: non N-maths, 1: N-maths)

- X_3 is a dummy variable for S-mathematics (0: non S-maths, 1: S-maths)

- X_4 is agreeableness (Likert scale 1–5, 1: strongly disagree, 5: strongly agree)

- X_5 is conscientiousness (Likert scale 1–5, 1: strongly disagree, 5: strongly agree)

- X_6 is extraversion (Likert scale 1–5, 1: strongly disagree, 5: strongly agree)

- X_7 is neuroticism (Likert scale 1–5, 1: strongly disagree, 5: strongly agree)

- X_8 is attendance per week (0 to 15)

- X_9 is study effort per week (hours)

- X_{10} is GPA from upper secondary school

- ε is stochastic error.

The 20-item version of the big-five inventory (BFI-20) applied in this paper is similar to the one developed by Engvik and Clausen (2011). The descriptive statistics presented in Tables 2a, 2b and 2c are applied for determining the coefficients in the regression model. To avoid multicollinearity, no variables for P-mathematics are included in the model. Students with a background in P-mathematics belong to the reference group. By presenting standardised Beta coefficients, we eliminate the differences in score values between MC and CR questions. The standardised values are independent of the unit's measurement. Hence, the rank and values of beta are comparable. Furthermore, with all VIF values below 2.0 there are no multicollinearity problems.

4.4   Findings

The linear regression model shows a significant link between mathematical background and performance in essay questions. It is strongest for the independent variable N-maths, but it is also substantial for students who chose S-maths at upper secondary school (beta = 0.205, p = 0.043). Furthermore, the independent variables gender, study effort and conscientiousness have a significant positive association with outcome of CR-based test.

Using MC scores as dependent variable shows another result. Neither mathematical background, study effort, nor conscientiousness are significantly correlated with success in the MC tests. This confirms H3, H4 and H7. For gender, however, there is a weak significant positive relationship. Adjusting for other factors, the gender effect in favour of males is a little bit stronger using the CR-format test compared to the MC-format test. The gender difference is rather small. Therefore, H2 is not confirmed. Independent of test methods, there is no association between performance in macroeconomics and the two dependent variables – attendance and GPA. H5 and H6 are rejected.

Table 5   Results from linear regression model

	Performance (dependent variable) multiple-choice test (MC)			Performance (dependent variable) essay question tests (CR)		
	Beta[1]	Sig.	VIF	Beta[1]	Sig.	VIF
(constant)		0.365			0.784	
Gender	0.171	0.074	1.426	0.205	0.024	1.426
N-maths	0.145	0.161	1.662	0.365	0.000	1.662
S-maths	0.021	0.843	1.783	0.206	0.042	1.783
Openness	−0.227	0.007	1.083	−0.167	0.034	1.083
Extraversion	0.077	0.410	1.372	0.140	0.114	1.372
Agreeableness	−0.071	0.415	1.210	−0.010	0.906	1.210
Conscientiousness	0.141	0.112	1.232	0.171	0.042	1.232
Emotional stability	0.022	0.826	1.603	−0.025	0.796	1.603
Attendance	0.137	0.110	1.145	−0.065	0.418	1.145
Study effort	0.002	0.979	1.217	0.195	0.020	1.217
GPA	0.047	0.581	1.136	0.069	0.391	1.136
	N = 145, adjusted R square = 0.090			N = 145, adjusted R square = 0.187		

Notes: [1]standardised coefficient beta.
    All VIF (variance inflation factor) values are below 2.0.

Openness is negatively correlated with achievement in macroeconomics (for both exam forms). There are no significant correlations for the three factors of agreeableness, extraversion and emotional stability.

In opposition to the CR-format test, there is almost no significant correlation between the independent variables and scores in the MC. Furthermore, the adjusted R square is substantially lower (adj. $R^2 = 0.09$) compared to the finding with CR as dependent variable (adj. $R^2 = 0.187$) with a value more than twice as high. The data indicate there is a substantial difference in factors associated with success in macroeconomics depending on whether the test method is MC or CR. Therefore, there is a variation in regard to which students have success in macroeconomics depending on the type of test.

5   Discussion

The findings for performance in essay questions are substantially in line with previous research. Findings confirm the result of Ballard and Johnson (2004) and Opstad (2018)

that mathematical skills are crucial factors for success in economics. Since N-mathematics is more theoretical and focuses on natural science, while S-mathematics is intended for business studies and social science, it is not obvious which choice of mathematics confers the greatest success in business studies. Opstad et al. (2017) reported mixed results. For quantitative-oriented courses like microeconomics and macroeconomics, N-maths students outperformed their counterparts, while S-maths students performed better in accounting and management. For business economics there was no difference in the positive impact on outcome by comparing N- and S-maths. Therefore, mathematical background from secondary upper school is a good predictor of performance in business courses. Despite this, by using MC as an exam format our data indicate no link between scores achieved and mathematical abilities, since the Beta coefficients are not significant. Hence, mathematical background is not a good predictor of MC scores. Students with good mathematical abilities achieve less by using MC instead of CR. For the independent variables, study effort, and conscientiousness (one of the personality traits), the result reveals the same finding. There is no significant correlation with MC-based test.

Numerous empirical studies have emphasised the association between study effort and performance in economics. The literature shows students' study time is an important and critical component for success (Rury and Carell, 2020).

Prior researchers are quite consistent in their analysis of the association between conscientiousness and academic success. Students described as hard working, achievement-focussed, disciplined, good at task planning, and who show goal-directed behaviour have appeared to demonstrate that these qualities are stable and a good predictor of academic success (Wang et al., 2019). Among personal characteristics, the trait of conscientiousness has the strongest impact on academic performance (Schneider and Preckel, 2017). Unlike the CR-formatted test, study effort and conscientiousness show no significant correlation with the MC-based test (H3 and H6).

Key factors to explain the performance in economics courses using essay tests are consistent with prior research. However, for the MC questions we do not find such a connection. The data are from the same student group, the same course, and the same exam day. Hard-working students get fewer rewards by substituting a CR-based exam with an MC-based exam.

In this study, factors explaining performance in the essay test do not apply for the MC test. Krieg and Uyar (2001) also suggest that the scores on MC tests do not verify overall student achievement in economics courses. Using a regression analysis, they find that the estimated coefficients for MC differ from CR. According to their study, gender and college GPA contribute more to performance in MC than CR. Our findings suggest there are substantial differences in the links between the dependent variables and performance in macroeconomics depending on the exam form, but in a different way than reported by Krieg and Uyar (2001). Unlike their findings, our research estimated a marginally higher male effect for CR questions than for MC questions.

Note that other key variables substantially contributed to explaining the performance in CR, but not in MC. Hence, one can question whether an MC exam actually measures students' knowledge and abilities in macroeconomics even when there is a significant correlation between the performance of the two test methods (see Table 4). For some students the exam format can be a critical factor for their performance in macroeconomics.

Prior research reports mixed results for the personality trait openness and performance. In line with the finding of De Fruyt and Mervielde (1996) and Busato et al. (2000), this study presents a negative correlation between openness and success in macroeconomics for both test methods (CR and MC). According to Vedel and Poropat (2017), openness is probably the most complicated and most debated factor to explain academic success. The factor includes intelligence, curiosity, creativity and artistic ability. There is no obvious reason why the personality trait of openness among business students has a low or even negative impact on performance. Lounsbury et al. (2009) suggest business students tend to be more conservative in their attitudes and behaviours. This is an element to take into consideration, but we need more research about the connection between openness and performance in economics.

According to previous publications, the mean correlations between academic performance and the three personality traits factors agreeableness, extraversion, and emotional stability are rather weak (O'Connor and Paunonen, 2007). Hence, it is no surprise that we did not find any significant connections between performance in macroeconomics and those personal characteristics. Students with those personality traits do not seem to be affected by the selection of the exam format discussed in this paper.

In general, previous studies report mostly a weak positive link between students' attendance and their performance in economics. Due to changes in technology, attendance is probably now a less important factor to predict achievement in economics (Büchele, 2020). Digital presentation and interactive learning through programs like Blackboard give students the opportunity to substitute their attendance at lectures with other study methods. This can explain why neither MC nor CR tests had a statistically significant correlation with students' attendance when controlling for gender, mathematical background, personality traits, GPA and study effort (H5 is not confirmed).

Proxies for students' abilities are their GPAs from secondary upper school. Therefore, it makes sense that there is a positive link between GPA and performance in economics courses. However, this study does not confirm this result either for CR or MC.

Findings in this study indicate that students' abilities, traits and backgrounds matter for achievement in the study of macroeconomics. Therefore, different students succeed in macroeconomics depending on the selection of MC or CR.

6   Limitations

The present research is from only one business school in Norway. Therefore, one must be careful about assuming the validity of these results in an international context. Hence, one cannot generalise the findings. The quality and the design of questions are important when one compares MC with CR (Chan and Kennedy, 2002; Harper, 2003). Maybe a different construction of MC would better capture the variation in students' mathematical skills, personality traits and effort. Therefore, more research on this issue is required to find out which students perform differently depending on the assignment format.

7   Conclusions

By using the same statistical methods and comparing statistical coefficients linked to performance for the same student group, in a cohort of business and economics students

in a Norwegian university, the results for MC questions and essay questions turned out very differently. For essay questions, the findings were mostly in line with prior research. However, for MC questions there were no positive significant connections between the dependent variables and MC scores except for gender and the personality trait of openness. Hence, it is unclear if the MC-format exam captures students' knowledge and skills in economics. This will have implications for the composition of students who achieve different performance in macroeconomics. MC format can easily measure recall and comprehension skills, but it is not as effective when it comes to capturing students' capacity for analysis, evaluation and creativity. Therefore, Douglas et al. (2012) recommend combining MC with other kinds of tests. It is important to keep applying tests that cover cognitive skills which strengthen students' abilities in writing with critical and analytical insight.

References

Aiken, L.R. (1982) 'Writing multiple-choice items to measure higher-order educational objectives', *Educational and Psychological Measurement*, Vol. 42, No. 3, pp.803–806 [online] http://doi.org/10.1007/bf03173468

Anderson, G., Benjamin, D. and Fuss, M.A. (1994) 'The determinants of success in university introductory economics courses', *The Journal of Economic Education*, Vol. 25, No. 2, pp.99–119 [online] http://doi.org/10.1080/00220485.1994.10844820.

Anderson, L.W., Krathwohl, D.R. and Airasian, P.W. (2001) *A Taxonomy for Learning, Teaching, and Assessing – A Revision of Bloom's Taxonomy of Educational Objectives*, Addison Wesley Longman, New York.

Arnold, I. and Straten, J.T. (2012) 'Motivation and math skills as determinants of first-year performance in economics', *The Journal of Economics Education*, Vol. 43, No. 1, pp.33–47 [online] http://doi.org/10.1080/00220485.2012.636709.

Arnold, I.J. and Rowaan, W. (2014) 'First-year study success in economics and econometrics: the role of gender, motivation, and math skills', *The Journal of Economic Education*, Vol. 45, No. 1, pp.25–35 [online] http://doi.org/10.1080/00220485.2014.859957.

Ballard, C. and Johnson, M.F. (2004) 'Basic math skills and performance in an introductory economics class', *Journal of Economic Education*, Vol. 35, No. 1, pp.3–23 [online] http://doi.org/10.3200/jece.35.1.3-23.

Bandura, A. (2010) 'Self-efficacy', *The Corsini Encyclopedia of Psychology*, pp.1–3, Wiley Online Library [online] http://doi.org/10.1002/9780470479216.corpsy0836.

Becker, W.E. and Johnston, C. (1999) 'The relationship between multiple choice and essay response questions in assessing economics understanding', *Economic Record*, Vol. 75, No. 4, pp.348–357 [online] http://doi.org/10.1111/j.1475-4932.1999.tb02571.x.

Bonesrønning, H. and Opstad, L. (2012) 'How much is students' college performance affected by quantity of study?', *International Review of Economics Education*, Vol. 11, No. 2, pp.46–63 [online] http://doi.org/10.1016/s1477-3880(15)30012-8.

Bonesrønning, H. and Opstad, L. (2015) 'Can student effort be manipulated? Does it matter?', *Applied Economics*, Vol. 47, No. 15, pp.1511–1524 [online] http://doi.org/10.1080/00036846.2014.997923.

Bridgeman, B. (1992) 'A comparison of quantitative questions in open-ended and multiple-choice formats', *Journal of Educational Measurement*, Vol. 29, No. 3, pp.253–271 [online] http://doi.org/10.1111/j.1745-3984.1992.tb00377.x.

Brookshire, R.G. and Palocsay, S.W. (2005) 'Factors contributing to the success of undergraduate business students in management science courses', *Decision Sciences Journal of Innovative Education*, Vol. 3, No. 1, pp.99–108 [online] http://doi.org/10.111/j.1540-4609.2005.0054.x.

Brown-Robertson, L.N., Ntembe, A. and Tawah, R. (2015) 'Evaluating the 'underserved student' success in economics principles courses', *Journal of Economics and Economic Education Research*, Vol. 16, No. 3, p.13.

Büchele, S. (2020) 'Evaluating the link between attendance and performance in higher education: the role of classroom engagement dimensions', *Assessment and Evaluation in Higher Education*, pp.1–19 [online] http://doi.org/10.1080/02602938.2020.1754330.

Buckles, S. and Siegfried, J.J. (2006) 'Using multiple-choice questions to evaluate in-depth learning of economics', *Journal of Economic Education*, Vol. 37, Winter, pp.48–57 [online] http://doi.org/10.3200/jece.37.1.48-57.

Buju, S. (2013) 'Personality profile of students with technical academic performance', *Procedia-Social and Behavioral Science*, Vol. 78, pp.56–60 [online] http://doi.org/10.1016/j.sbspro.2013.04.250.

Busato, V.V., Prins, F.J., Elshout, J.J. and Hamaker, C. (2000) 'Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education', *Personality and Individual Differences*, Vol. 29, No. 6, pp.1057–1068.

Cannonier, C. and Smith, K. (2019) 'Do crib sheets improve student performance on tests? Evidence from principles of economics', *International Review of Economics Education*, Vol. 30, p.100147 [online] http://doi.org/10.1016/j.iree.2018.08.003.

Chan, N. and Kennedy, P.E. (2002) 'Are multiple-choice exams easier for economics students? A comparison of multiple-choice and 'equivalent' constructed-response exam questions', *Southern Economic Journal*, pp.957–971 [online] http://doi.org/10.2307/1061503.

Chowdhury, M. (2006) 'Students' personality traits and academic performance: a five-factor model perspective', *College Quarterly*, Vol. 9, No. 3, p.3.

Cohn, E. and Johnson, E. (2006) 'Class attendance and performance in principles of economics', *Education Economics*, Vol. 14, No. 2, pp.211–233.

Costa, P.T. and McCrae, R.R. (1992) *Revised NEO Personality Inventory and NEO Five Factor Inventory Professional Manual*, Psychological Assessment Resources, Odessa, FL [online] http://doi.org/10.1037/t07553-000.

De Fruyt, F. and Mervielde, I. (1996) 'Personality and interests as predictors of educational streaming and achievement', *European Journal of Personality*, Vol. 10, No. 5, pp.405–425 [online] http://doi.org/10.1002/(sici)1099-0984(199612)10:5.

Douglas, M., Wilson, J. and Ennis, S. (2012) 'Multiple-choice questions tests: a convenient, flexible and effective learning tool? A case study', *Innovations in Education and Teaching International*, Vol. 49, pp.111–121 [online] http://doi.org/10.1080/14703297.2012.677596.

Dwyer, C.P., Hogan, M.J. and Stewart, I. (2014) 'An integrated critical thinking framework for the 21st century', *Thinking Skills and Creativity*, Vol. 12, pp.43–52 [online] http://doi.org/10.1016/j.tsc.2013.12.004.

Engvik, H. and Clausen, S.E. (2011) 'Norsk kortversjon av big five inventory (BFI-20)', *Tidsskrift for Norsk Psykologforening*, in Norwegian, Vol. 48, No. 9, pp.869–872.

Fallan, L. and Opstad, L. (2014) 'Beyond gender performance in accounting: does personality distinction matter?', *Accounting Education*, Vol. 23, No. 4, pp.343–361 [online] http://doi.org/10.1080/09639284.2014.930693.

Fallan, L. and Opstad, L. (2016) 'Student self-efficacy and gender-personality interactions', *International Journal of Higher Education*, Vol. 5, No. 3, pp.32–44 [online] http://doi.org/10.5430/ijhe.v5n3p32.

Guillaume, D.W. and Khachikian, C.S. (2011) 'The effect of time-on-task on student grades and grade expectations', *Assessment and Evaluation in Higher Education*, Vol. 36, No. 3, pp.251–261 [online] http://doi.org/10.1080/02602930903311708.

Harper, R. (2003) 'Multiple-choice questions – a reprieve', *Bioscience Education*, Vol. 2, No. 1, pp.1–6, DOI: 10.3108/beej.2003.02000007.

Hickson, S., Reed, W.R. and Sander, N. (2012).'Estimating the effect on grades of using multiple choice versus constructive-response questions: data from the classroom', *Educational Assessment*, Vol. 17, No. 4, pp.200–213 [online] http://doi.org/10.1080/10627197.2012. 735915.

Horn, P.M. and Jansen, A.I. (2009) 'Tutorial classes – why bother? An investigation into the impact of tutorials on the performance of economics students', *South African Journal of Economics*, Vol. 77, No. 1, pp.179–189 [online] http://doi.org/10.1111/j.1813-6982. 2009.01194.x.

Ibrahim, N.S., Yusof, C.N.S.H., Razak, N.A. and Norshahidi, N.D. (2014) 'A meta-analysis of the relationship between big five personality traits and students' academic achievement', *ICSSR-e-Journal of Social Science Research*, Vol. 1, No. 2, pp.15–22.

Johnson, M., Robson, D. and Taengnoi, S. (2014) 'A meta-analysis of the gender gap in performance in collegiate economics courses', *Review of Social Economy*, Vol. 72, No. 4, pp.436–459 [online] http://doi.org/10.1080/00346764.2014.958902.

Kassis, M.M. and Boldt, D.J. (2020) 'Factor impacting student success in introductory economics courses', *Journal for Economic Educators*, Vol. 20, No. 1, pp.41–63.

Krieg, R.G. and Uyar, B. (2001) 'Student performance in business and economics statistics: does exam structure matter?', *Journal of Economics and Finance*, Vol. 25, No. 2, pp.229–241 [online] http://doi.org/10.1007/bf02744525.

Kuechler, W.L. and Simkin, M.G. (2010) 'Why is performance on multiple-choice tests and constructed-response tests not more closely related? Theory and an empirical test', *Decision Sciences Journal of Innovative Education*, Vol. 8, No. 1, pp.55–73 [online] http://doi.org/ 10.1111/j.1540-4609.2009.00243.x.

Lounsbury, J.W., Smith, R.M., Levy, J.J., Leong, F.T. and Gibson, L.W. (2009) 'Personality characteristics of business majors as defined by the big five and narrow personality trait', *Journal of Education for Business*, Vol. 84, No. 4, pp.200–205 [online] http://doi.org/ 10.3200/JOEB.84.4.200-205.

Mallik, G. and Lodewijks, L. (2010) 'Student performance in a large first year economics subject: Which variables are significant?', *Economic Paper*, Vol. 29, No. 1, pp.80–86 [online] http://doi.org/10.1111/j.1759-3441.2010.00051.x.

Matthews, G. and Oddy, K. (1993) 'Recovery of major personality dimensions from trait adjective data', *Personality and Individual Differences*, Vol. 15, No. 4, pp.419–431 [online] http://doi.org/10.1016/0191-8869(93)90070-j.

O'Connor, M.C. and Paunonen, S.V. (2007) 'Big five personality predictors of post-secondary academic performance', *Personality and Individual Differences*, Vol. 43, pp.971–990 [online] http://doi.org/10.1016/j.paid.2007.03.017.

Opstad, L. (2018) 'Success in business studies and mathematical background: the case of Norway', *Journal of Applied Research in Higher Education*, Vol. 10, No. 3, pp.399–408, DOI: 0.1016/ S1477-3880(15)30059-1.

Opstad, L. and Fallan, L. (2010) 'Student performance in principles of macroeconomics: the importance of gender and personality type', *International Review of Economics Education*, Vol. 9, No. 1, pp.76–92 [online] http://doi.org/10.1016/s1477-3880(15)30059-1.

Opstad, L., Bonesrønning, H. and Fallan, L. (2017) 'Tar vi opp de rette studentene ved økonomisk-administrative studier?', *Samfunnsøkonomen*, in Norwegian, Vol. 31, No. 1, pp.21–29.

Park, K.H. and Kerr, P.M. (1990) 'Determinants of academic performance: a multinomial logit approach', *The Journal of Economic Education*, Vol. 21, No. 2, pp.101–111.

Parker, K. (2006) 'The effect of student characteristics on achievement in introductory microeconomics in South Africa', *South African Journal of Economics*, Vol. 74, No. 1, pp.137–149 [online] http://doi.org/10.1111/j.1813-6982.2006.00054.x.

Rodgers, J.R. (2001) 'A panel-data study of the effect of student attendance on university performance', *Australian Journal of Education*, Vol. 45, No. 3, pp.284–295.

Rothstein, M.G., Paunonen, S.V., Rush, J.C. and King, G.A. (1994) 'Personality and cognitive ability predictors of performance in graduate business school', *Journal of Educational Psychology*, Vol. 86, No. 4, p.516 [online] http://doi.org/10.1037/0022-0663.86.4.516.

Rury, D. and Carrell, S. (2020) *Knowing What it Takes: The Effect of Information about Returns to Studying on Study Effort and Achievement*, July [online] http://conference.iza.org/conference_files/edu_2020/rury_d29540.pdf (accessed 15 April 2021).

Sansani, S. and Rahamim, A. (2019) 'Available study time and undergraduate student exam performance', *Journal of Applied Research in Higher Education*, Vol. 11, No. 1, pp.20–35 [online] http://doi.org/10.1108/JARHE-12-2017-0158.

Schneider, M. and Preckel, F. (2017) 'Variables associated with achievement in higher education: a systematic review of meta-analyses', *Psychological Bulletin*, Vol. 143, No. 6, p.565 [online] http://doi.org/10.1037/bul0000098.

Simkin, M.G. and Kuechler, W.L. (2005) 'Multiple-choice tests and student understanding: what is the connection?', *Decision Sciences Journal of Innovative Education*, Vol. 3, No. 1, pp.73–98 [online] http://doi.org/10.1111/j.1540-4609.2005.00053.x.

Stinebrickner, R. and Stinebrickner, T.R. (2008) 'The causal effect of studying on academic performance', *The BE Journal of Economic Analysis and Policy*, Vol. 8, No. 1, pp.1–53 [online] http://doi.org/10.2202/1935-1682.1868.

Swope, K.J. and Schmitt, P.M. (2006) 'The performance of economics graduates over the entire curriculum: the determinants of success', *Journal of Economic Education*, Vol. 37, pp.387–394 [online] http://doi.org/10.3200/jece.37.4.387-394.

Trapman, S., Hell, B., Hirn, J.W. and Schuler, H. (2007) 'Meta-analysis of the relationship between the big five and academic success at university', *Journal of Psychology*, Vol. 215, No. 2, pp.132–151 [online] http://doi.org/10.1037/e518532013-271.

Vedel, A. and Poropat, A.E. (2017) 'Personality and academic performance', *Encyclopedia of Personality and Individual Differences*, pp.1–9 [online] http://doi.org/10.1007/978-3-319-28099-8_989-1.

Walstad, W.B. and Becker, W.E. (1994) 'Achievement differences on multiple-choice and essay tests in economics', *The American Economic Review*, Vol. 84, No. 2, pp.193–196 [online] http://doi.org/10.2307/1182910.

Wang, S., Zhao, Y., Li, J., Wang, X., Luo, K. and Gong, Q. (2019) 'Brain structure links trait conscientiousness to academic performance', *Scientific Reports*, Vol. 9, No. 1, pp.1–12 [online] http://doi.org/10.1038/s41598-019-48704-1.

Watts, M. (1987) 'Student gender and school district differences affecting the stock and flow of economic knowledge', *Review of Economics and Statistics*, Vol. 69, pp.561–566 [online] http://doi.org/10.2307/1925550.