

Testing the effectiveness of in-semester assessment in Econ 101

NZAE conference 1-3 July 2009

We compare the performance of course-work (take-home assignments, on-line quizzes and invigilated tests) against the final examination in Econ 101 over 16 semesters (2001-8). When course-work is less comprehensive and less intensive than the final examination, and so less challenging, completion of coursework acts more as a signal of a student's participation than of their ability per se. Open-book assessment (assignments and on-line quizzes) and tests with only multiple-choice questions all proved limited as predictors of success in the subsequent examination, though more robust as predictors of examination failure. We found tests that required written answers were better indicators of examination performance than tests with multiple-choice only. Coursework that does not duplicate the scope and intensity of the final examination is not a substitute for a final examination, but is a complement which engages students throughout the semester. Awarding marks for coursework provides an incentive to study and reduces the crowding out of assessment by the demands of other courses.

Gillis Maclean and Paul McKeown, Lincoln University

JEL classification: A22 - Economics Education and Teaching of Economics – Undergraduate

Note: this is a preliminary draft and intended for use only at this NZAE conference.

1. Introduction

As teachers we assess the performance of our students as a matter of course. From time to time we need to examine ourselves, to take a step back and assess our assessment, focusing not on students' performance but on the performance of our assessment measures. How effective is our assessment? Here we investigate how useful the forms of in-semester assessment used in Econ 101 are for the teacher and for the student. First, how well does in-semester assessment predict the likelihood of passing the final exam? This is critical information for students in their study before the exam. It also helps teachers in identify where intervention may be needed during the semester, to help struggling students, and after the exam when considering aegrotats. Is passing or failing an in-semester assessment better indicator of exam performance? Second, how well does in-semester assessment provide incentives that keep students engaged in learning throughout the semester?

Other questions include: What is the most useful balance between formative and summative assessment? What is the most useful balance between multiple-choice and written answers for in-semester assessment? What is the most useful balance between in-semester assessment and the final exam? These questions can be answered only in the context of the course in question; we present our findings for Econ 101.

Lincoln University's Econ 101 (Principles of Economics) was the introductory economics course, covering both micro and macroeconomics as part of the compulsory core for the BCM degree. It was the pre-requisite for the second introductory combined micro/macro course, Econ 104 (Foundations of Economic Policy). Despite some personnel changes, Econ 101 was offered in substantially the same form over the years 2001-08, with a consistent teaching culture and intended level of difficulty; both authors were involved throughout this time. This paper outlines our initial exploration of our database from these 16 semesters. The results presented here are from a low resolution analysis of the data, with our main focus being on how well passes/fails (and participation) of in-semester assessment predict passes/fails in the final exam. This is the first run of a work-in-progress and we intend further analysis of the data used here (such as significance testing). We will also be adding more recent and more detailed data (e.g. full web-site access logs and the results of web surveys) to analyse student study habits and perceptions.

Section 2 describes the forms of assessment, the data series and our approach. Section 3 describes our results and Section 4 concludes.

2. Data and Method

In Econ 101 we use 2 types of questions for all assessment: multiple-choice (hereafter MC) questions and short-answer questions requiring written answers including definitions, descriptions and explanations, drawing and interpreting diagrams, and calculations (hereafter SA). There were 4 forms of in-semester assessment: (1) take-home assignments, (2) on-line quizzes, and invigilated tests with (3) multiple-choice only and with (4) multiple-choice plus written answers. We compare the performance of each form of assessment against the final exam, which is traditionally accepted as the most complete and accurate form of assessment and despite changes in technology remains the ultimate form of assessment for undergraduate courses. For example, by regulation, all Lincoln University undergraduate courses must have at least 40% weight on the final exam.

We break the data into 3 time periods: 2001-03, 2004-06 and 2007-08 (see Appendix 1 for assessment details). The final exam format (29% MC and 71% SA) and weighting (50% of final mark) remained constant over all three periods. The first 2 periods share the same curriculum but have different in-semester assessment (coursework). In the first (2001-03), coursework consisted of 5 take-home assignments with SA questions only (worth 30%) and 2 invigilated MC-only tests (each worth 10%). Assignments proved unsuitable and were not used after 2003, and the short-answer questions were moved from the assignments into the invigilated tests. In the second period (2004-06), coursework consisted of just 2 tests that each contained a combination of MC and SA questions. The second test had more written questions and more weighting (with test 1 worth 20% and test 2 worth 30% of the overall course mark). In the third period (2007-08), the course was restructured with changes to curriculum, teaching and overall pass rates, and so is taken as a distinct population. The balance of the tests was further adjusted with more SA and less MC questions, and on-line MC quizzes were introduced (2008).

Assessment outcomes depend on many variables. Some we can control, such as clarity of learning objectives, the fit of assessment to the material being assessed, and consistency of assessment over time; hence the division of our data into 3 populations. Others we cannot control, such as students' ability and motivation. However, we can use assessment completion rates as a proxy for the level of a student's engagement in the course. We take the attempt of all relevant assessment as an indicator that a student is actively engaged in their study over the semester, at least at some minimal level. (A future project is to analyse whether the marks a student has already gained appear to change their motivation in the next assessment.)

To present the best case for the in-semester assessment, we consider only students who attempted all relevant assessment, and thus showed some active engagement over the whole semester. Students who did not attempt all relevant assessment (regardless of any aegrotat) are excluded from the population because they were not actively engaged for, at least, part of the semester. We separate these two groups because of their different behaviours – and very different assessment outcomes. We know that disengaged students are likely to fail; our question is, what does engagement in in-semester assessment tell us about students' subsequent exam performance? In this best-case scenario, assessment pass rates are not dragged down by those who fail because of non-completion. We have also thereby excluded all students who applied for an aegrotat for any piece of assessment, although by presumption these students were actively engaged in study despite not being able to attempt everything. These are excluded because an aegrotat mark can only be estimated after the fact.

To analyse the performance of the two types of questions used in the tests (MC and SA), we take the simple total of the marks for each question type over the two tests and compare them to the relevant section's mark in the final exam. However, for students the important issue is the aggregate weighted coursework mark. To capture the effect of coursework, aggregate test marks are recorded with relative coursework weightings rather than as simple totals (see Appendix 1).

We then look at participation rates and compare the outcomes for those students who did not complete all relevant assessment. Whether a student actually attempts assessment is itself a key indicator. A critical role in course failures is student drop-out during the semester – the DNS (Did Not Sit).

2.1. Assessment overview

Each assignment (18 marks) was made up of a several short-answer questions similar in format to the final exam. Five assignments were submitted per semester, each worth 6%, for a total of 30% of the final mark. Students had 2 weeks to hand in each assignment; given that an assignment could be completed in several hours, for all practical purposes the time was unlimited. The take-home assignment was not only open-book, but allowed students to work together and ask for help from tutors and lecturers (with the usual restrictions on copying). Under these conditions no student applying effort is expected to fail. Assignments were seen as providing mostly formative assessment; the purpose was to get students to learn by practising.

Over the entire time (2001-08) there were 2 invigilated tests per semester, each test assessing 4 weeks work. The test format evolved over the 3 periods. Initially (2001-3), as we set written assignments,

both tests were multiple-choice only (each with 30 MC in 50 minutes and each worth 10% of the final mark). In 2004, when assignments were dropped from the course, we added written answers to the tests so the tests became a combination of MC and SA. For 2004-06, test 1 had 30 MC and 10 marks SA in 90 minutes; test 2 was more challenging with 30 MC and 20 marks SA in 120 minutes. The test weightings increased to 20% and 30% respectively (giving SA questions more weight in the overall course mark). In the third period, the MC were reduced to 26 and the SA increased to 14 and 24 respectively, with the same time allowance and weighting for final mark. Lastly in semester 2 2008 with the introduction of on-line MC quizzes, the balance in the tests was further shifted from MC to SA, and the weighting reduced by 5% per test to accommodate the quiz marks.

As well as 2 tests and the exam, we wanted an extra assessment format which created frequent engagement by students, provided rapid feedback to students and teachers, and carried a grade as an incentive for students to participate. On-line quizzes (via Moodle) were introduced in 2008. There were 12 weekly quizzes and each consisted of 10 MC questions selected randomly from a bank of around 30 questions. In 2008 semester 1 we piloted the quizzes without awarding marks for them. In Semester 2, quizzes contributed 10% to the overall course mark. Students can make unlimited attempts for any quiz and the best mark is recorded, with the average of the best 10 quizzes being used as the overall quiz mark.¹ We carry out analysis of the quiz mark in the same way as the other forms of assessment for this one semester.

Although the final exam format, coverage and time (190 minutes) remained constant over the 8 years, the exam is not an “exogenous” variable or fixed standard against which we can measure in-semester assessment. Exam pass rates and means are liable to change over time, whether intentionally or not. Table 1 compares the 3 periods (excluding all DNS and aegrotats).

Table 1: Overall course and final exam results 2001-08

period	Econ 101 course		final exam	
	mean mark	pass rate	mean	pass rate
2001-03	59.9	73.0%	51.85	56.0%
2004-06	55.6	68.0%	53.77	62.0
2007-08	59.41	74.5%	60.82	72.0%

In the first period, to compensate for the inflation of marks with the 30% weight for the assignments, the final exam needed to be more challenging. In the second period, the overall course mean and pass rate were reduced, even though the removal of the assignment marks allowed an easier exam (with the

¹ A future project is to analyse how the change from zero to 10% of final marks changed student behaviour.

exam mean and pass rate actually increasing). In the third period, the course mean and pass rate were similar to the first period. More challenging in-semester tests in the third period had also allowed a further easing in the exam. The exam pass rate in period 3 was significantly higher than the pass rate for the course overall, a balance we think desirable (for reasons discussed below).

3. Results

Repeating students are counted in all semesters that they appear. All marks are converted to percentages, and a mark of 50% or greater is considered to be a pass. (Note totals in tables may not tally due to rounding.)

For each period we present a scatter diagram of marks for the assessment in question against the exam, where each point represents one student. If the two pieces of assessment were of equal difficulty, we would expect marks to be equally distributed along both sides of the $y = x$ line (with random variation). The more points there are above the $y = x$ line, the greater the number of students who are scoring better in the exam than the in-semester assessment. For each period, the “A” suffixed Tables present the percentage of students above the $y = x$ line, along with the pass rates and the means for the 2 pieces.

The “B” suffixed tables show the exam pass/fail conditional on the prior in-semester result, which is the telling data for predicting exam performance from in-semester assessment. Tables C (in Appendix 2) show the four possible pass/fail combinations by percentage of total students, corresponding to the 4 quadrants on the scatter diagram (mark either less than 50 or at least 50).

3.1. 2001 – 2003

Assignments (2001-03)

“Assignments” refers to the total mark for all 5 assignments (as a percentage), and “Exam” refers to the total mark for the final exam (as a percentage). The population is all students ($n_A = 1524$) over 6 semesters (2001-03 inclusive) who completed all assignments and sat the final exam.

Fig.1

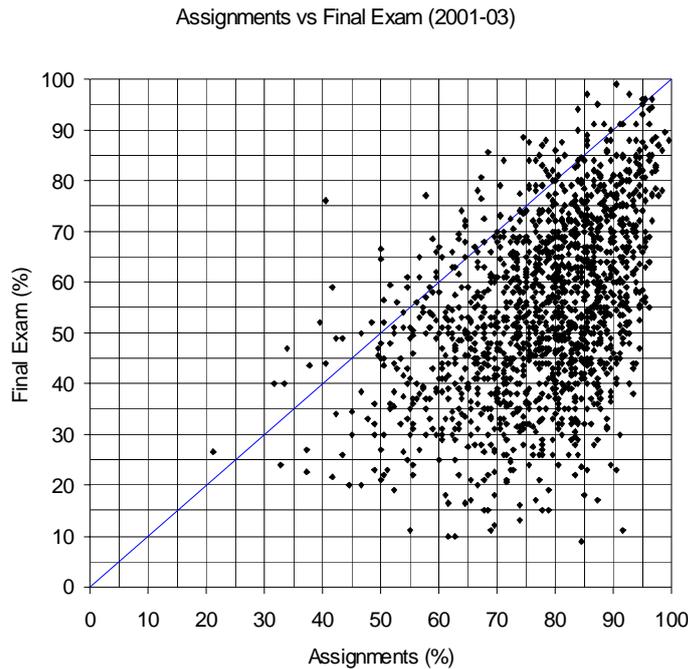


Table 2A: assignments and final exam 2001-03

	mean mark	pass rate
assignments	78.0	98.0%
exam	54.7	61.3%
% students scoring more in exam than assignments = 6.5%		

Assignment marks are crowded around the mean of 78. Only 6.5% of the population lies above the $y = x$ line; very few students did better in the final exam than in the 5 assignments. It is very much easier to gain marks in the assignments (mean mark 78.0) than in the exam (mean 54.7). Given the take-home conditions, students *should* pass the assignments, and indeed 98% of those completing did so. However, 96% of those who failed the final exam had passed the total assignments during the semester.

Table 2B: assignments and final exam 2001-03

% of students by row	pass exam	fail exam	row total
pass assignments	62.2%	37.8%	100%
fail assignments	16.1%	83.9%	100%

The first row in Table 2B shows the percentage of students who passed the assignments and went on to sit and pass or fail the exam; the second row shows the percentage of students who completed but failed the assignments and went on to pass or fail the exam.

A pass in the assignments was not a good indicator of performance in the final exam: only 62.2% of students who passed the assignments went on to pass the exam, and this rate is no better than for students who simply completed all assignments (61.3%). The fact that a student passed the assignments gives no more information than the fact that the student merely completed them. Because take-home assignments are so much less difficult than the exam, an assignment pass tells us only that the student is still engaged – or at least, has not dropped out. However, competing but failing the assignments is a good indicator of failure in the exam; 83.9% of those who completed but failed the assignments then failed the exam.

Multiple-choice Tests (2001-03)

“Tests” refers to the total mark for both tests, and “Exam” refers to the total mark for the final exam. Because in this period the tests had equal weighting, the test total coincides with the weighted test mark. The population is all students over 6 semesters (2001-2003 inclusive) who completed both tests and sat the final exam ($n_T = 1782$); 17% more students sat both tests and the exam than completed all 5 assignments and sat the exam. Note that the expected test mark with random choice on the MC is 20%; only 1 student (of 1782) scored below 20% in the tests. This explains the lack of data points below Test MC = 20.

Fig.2

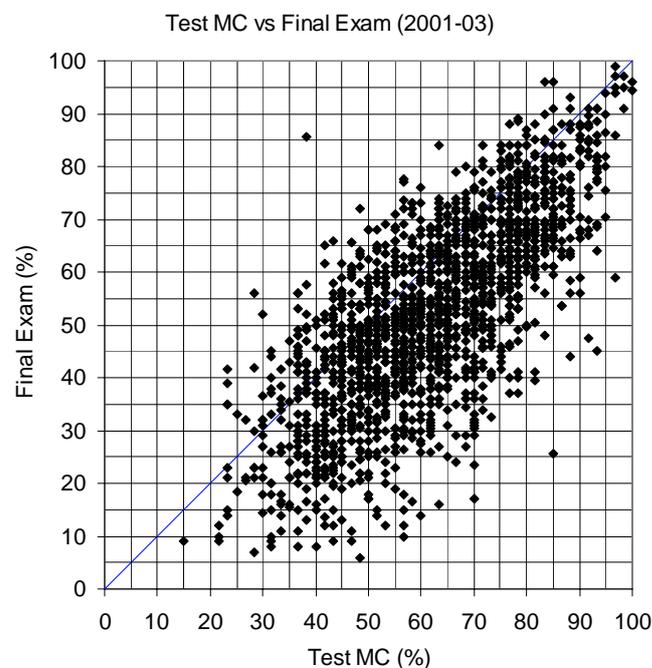


Table 3A: MC-only tests and total final exam 2001-03

	mean mark	pass rate
tests (MC-only)	61.8	76.8%
exam	52.1	55.4%
% students scoring more in exam than MC tests = 20.6%		

Compared to the assignments, test marks are spread more evenly along the $y = x$ line and less concentrated at the high end. MC tests are related more closely to the exam than the assignments, but the bulk of students are still below the $y = x$ line, with 20.6% of students scoring higher in the exam than the tests (which is not as extreme as the 6.5% figure for assignments). The multiple-choice tests were more difficult than the assignments; only 76.8% passed the MC tests (98% for assignments). As expected, the mean mark and pass rate for the less comprehensive MC tests (61.8 and 76.8%) are clearly higher than the mean and pass rate for the final exam (52.1 and 55.4%).

It may appear anomalous that, when the tests are harder than the assignments, the exam pass rate for students completing both tests (55.4%) is lower than for students completing all 5 assignments (61.3%); but we are comparing two different populations. The larger population (completed 2 tests versus completed all 5 assignments) now includes more partially engaged students who sat the tests, but did not complete the assignments, and then went on to fail the exam.

Table 3B: MC only tests and total final exam 2001-03

% of students by row	pass exam	fail exam	row total
pass MC tests	67.6%	32.4%	100%
fail MC tests	15.2%	84.8%	100%

The rows in Table 3B show the percentage of students who passed/failed the MC-only tests and went on to pass or fail the exam. The results are very similar to assignment performance for students completing the respective assessment; 67.6% of students who passed the tests went on to pass the exam (62.2% for assignments), and 84.8% of those who failed the tests then failed the exam (83.9% for assignments). However, the exam pass rate for those students who actually passed both tests (67.6%) is 20% better than the exam pass rate for students who simply completed both tests (55.4%), which is not the case for the assignments. Unlike assignment performance, that a student passes (rather than merely sits) the tests does give some information about subsequent exam performance.

Of the students who sat both tests, 55.8% of those who failed the final exam had previously passed the MC tests during the semester. This is very different from the corresponding assignment results: 96% of those who failed the final exam had passed the assignments.

3.2. 2004 – 2006

2004-06 test multiple-choice vs. exam multiple-choice

Here we take only the MC part of the 2 tests (unweighted total) and the MC part of the exam (29% of final exam). The population includes only those students who completed both tests and the final exam 2004-06 ($n = 2010$).

Fig.3

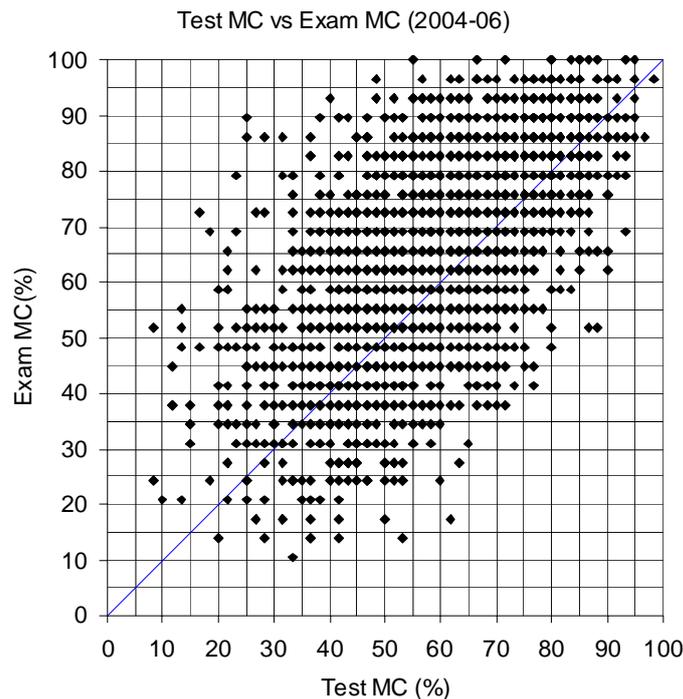


Table 4A: test MC & exam MC 2004-06

	mean mark	pass rate
test MC	58.4	71.6%
exam MC	64.6	77.9%
% students scoring more in exam MC than tests MC = 64.3%		

Note that given random choice one can expect 20% in the MC, both series should start from 20%; no students managed to get below this for either assessment. Compared to the assignments/exam and test/exam scatters for 2001-03, the majority of the population (64.3%) lies above the $y = x$ line; it is slightly easier to gain the same percentage for the MC in the exam (mean mark 64.6) than in the combined tests (mean mark 58.4). The pass rate for the MC part of the exam (77.9%) is slightly higher than the pass rate for the MC part of the combined tests (71.6%), indicating that the exam MC was generally easier than the test MC.

Although the test MC questions themselves were not more difficult than in the first period, the pass rate for the test MC is slightly reduced compared to 2001-03 (from 76.8% to 71.6%); this can be expected if the overall test is now longer and more difficult with the addition of short-answer questions.

The rows in Table 4B show the percentage of students who passed/failed the multiple-choice part of the tests and went on to pass or fail the multiple-choice part of the exam.

Table 4B: test MC & exam MC 2004-06

% of students by row	pass exam MC	fail exam MC	row total
pass test MC	88.2%	11.8%	100%
fail test MC	51.9%	48.1%	100%

Of those who passed the test MC, 88.2% went on to pass the exam MC. However, those who failed the test MC had about a 50/50 chance of passing the subsequent exam MC (51.9% vs. 48.1%). Hence, passing the harder test MC indicates a higher chance of passing the exam MC, whereas failing the harder tests doesn't necessarily mean that a student will fail the exam. This is consistent with students finding the MC part of the exam easier than the test MC. The random guessing component of the multi-choice tests may also lead to more variation in the scores of students who are know fewer answers (e.g. a good student will be doing less guessing and there will be less variation in their scores as a result); it may be harder to get 50% of 60 questions over 2 tests than 50% of 29 questions in the exam.

2004-06 test short-answer vs. exam short-answer

Here we take only the SA part of the 2 tests (unweighted total) and the SA part of the exam (71% of final exam). The population includes only those students who completed both tests and the final exam (n = 2010).

Fig.4.

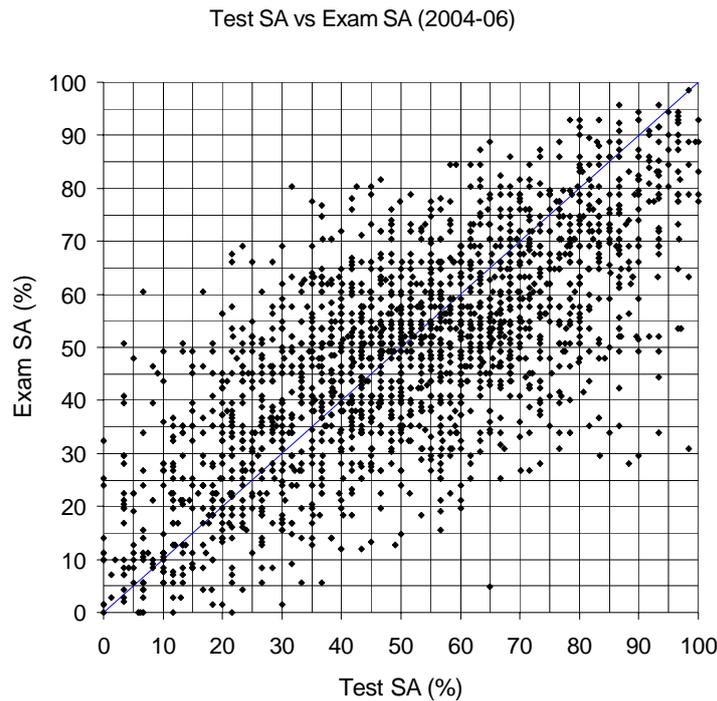


Table 5A: test SA & exam SA 2004-06

	mean mark	pass rate
test SA	51.4	53.3%
exam SA	49.1	50.6%
% students scoring more in exam SA than tests SA = 46.2%		

Compared to 2001-03, there are changes in the test/exam relationship because both pieces changed: the exam was made easier (with increases in both mean and pass rate) and the test was made more difficult by adding SA (with the SA part of the tests having a lower mean and pass rate than the MC tests of the first period). The combined effect is to shift the x/y scatter points upward compared to the corresponding test/exam scatter for 2001-03.

As expected, in this period the pass rate for the SA part of the exam (50.6%) is much lower than the MC part of the exam (77.9%) in the second period. The SA/SA points start at (0, 0) unlike the MC/MC which start at (20, 20). The two SA assessments are almost equally spread above and below the $y = x$ line, with 46.2% of the population above; SA marks are just easier to gain in the tests than the exam, with means of 51.4 and 49.1 respectively. The pass rate for the SA part of the tests (53.3%) is only just higher than the pass rate for the SA part of the exam (50.6%), which suggest equivalent levels of difficulty (although there are fewer SA questions in the tests than in the exam).

The rows in Table 5B show the percentage of students who passed (failed) the short-answer part of the tests and went on to pass or fail the short-answer part of the exam.

Table 5B: test SA & exam SA 2004-06

% of students by row	pass exam SA	fail exam SA	row total
pass test SA	72.8%	27.2%	100%
fail test SA	25.2%	74.8%	100%

Note the top-left/bottom-right diagonal: the test SA passes result in exam SA passes at about the same rate as the test SA fails result in exam SA fails. Of those who passed the test SA, 72.8% went on to pass the exam SA; this is lower than the corresponding rate for the MC (88.2%). However, of those who failed the test SA, only 25.2% went on to pass the subsequent exam SA; this is half the corresponding rate for the multiple-choice (51.9%). Overall, the test SA questions seem to be of similar difficulty to the exam SA questions. If a student fails the SA in the tests then they are most likely to fail the exam SA. Similarly, if a student passes the SA in the tests then they are most likely to pass the exam SA. Compared to the test MC, passing the test SA is not quite as good as a predictor of success in the corresponding part of the exam, but failing the test SA is a much better indicator of failure in the corresponding part of the exam.

Aggregate Tests 2004-06

“Aggregate tests” refer to the weighted coursework mark for both tests (combined MC and SA). The aggregated (weighted) test mark puts more weight on written answers for the SA questions and less on the MC questions by putting more weight on test 2 which has a larger SA component. The population is all students who sat both tests and the final exam ($n = 2010$).

Fig.5.

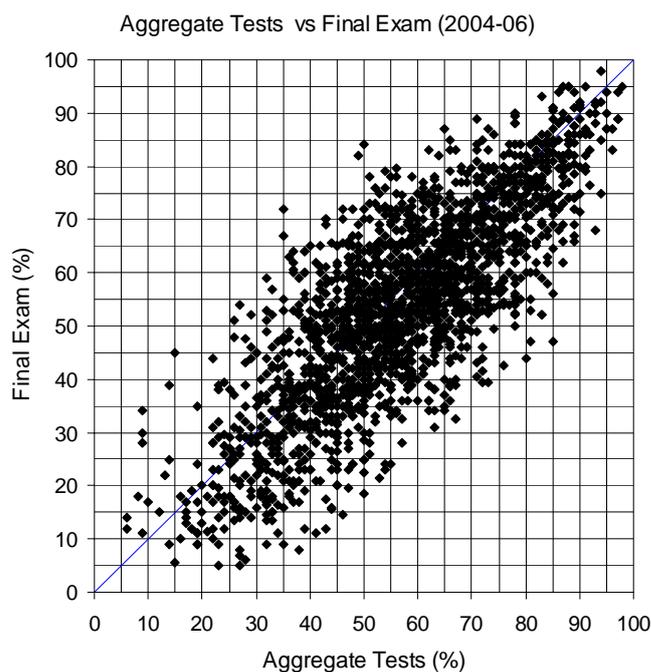


Table 6A: aggregate tests 2004-06

	mean mark	pass rate
aggregate test	56.7	67.0%
exam	53.6	62.1%
% students scoring more in exam than aggregate tests = 36.1%		

The two assessments are fairly equally spread above and below the $y = x$ line, with 36.1% of the population above, with the aggregate tests slightly easier than the final exam. As expected, inclusion of written questions increased the difficulty of tests compared to MC-only tests, and the aggregate test mean mark and pass rate fell from 61.8 and 76.8% (2001-03) to 56.7 and 67.0% respectively. At the same time, there was an increase in the exam pass rate for engaged students (merely completing, as opposed to actually passing both tests), from 55.4% in 2001-03 to 62.1% in 2004-06. This indicates that the exam became easier, but the positive impact of this on the overall pass rate was outweighed by the negative impact caused by the removal of the 30% of easily gained marks for assignments.

The rows in Table 6B show (for combined MC and SA) the percentage of students who passed (failed) the weighted tests and went on to pass or fail the final exam.

Table 6B: aggregate tests 2004-06

% of students by row	pass exam	fail exam	row total
pass tests	81.0%	19.0%	100%
fail tests	23.8%	76.2%	100%

Here the tests are more difficult and correspond (especially after weighting) more closely to the exam format than in 2001-03; the aggregate test mark is a thus better indicator of later exam performance when SA questions are included. The combined effect of more difficult tests and an easing in the exam is an increase in the pass exam/pass test ratio (from 67.6% to 81.0%) and a decrease in the fail exam/fail test ratio (from 84.8% to 76.2%). Those who pass the harder tests are more likely to pass the easier exam, and vice versa.

3.3. 2007 – 2008

2007-08 test multiple-choice vs. exam multiple-choice

Here we take only the MC part of the 2 tests (unweighted total) and the MC part of the exam; the number of MC questions in each test is reduced from the second period (2004-06). The population includes only those students who completed both tests and the final exam ($n = 775$).

Fig.6.

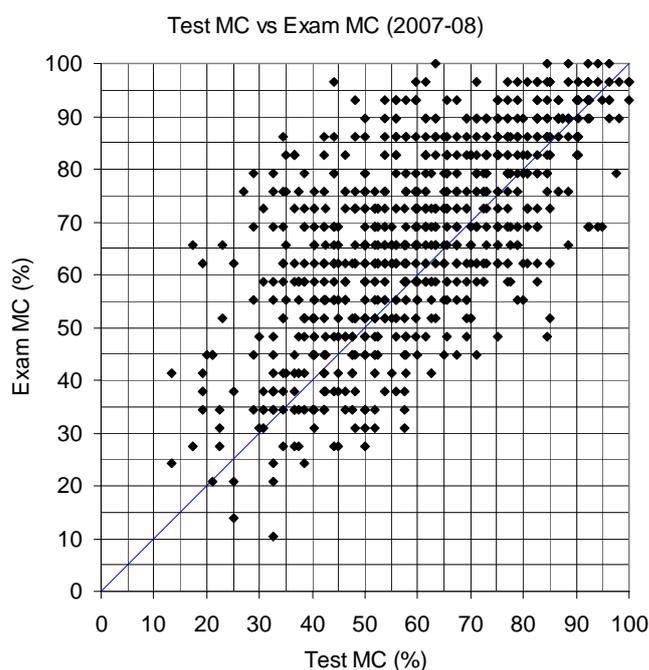


Table 7A: test MC & exam MC 2007-08

	mean mark	pass rate
test MC	59.8	71.2%
exam MC	67.2	82.2%
% students scoring more in exam MC than tests MC = 36.1%		

The spread is very similar to that of 2004-06; note again the spread starts at around (20, 20). Here, 69.0% of the population lies above the $y = x$ line; this is similar to the second period, i.e. it is easier to gain the same percentage for the MC in the exam (mean mark 67.2) than in the combined tests (mean mark 59.8). Despite a reduction in the number of multiple-choice questions in the test, the pass rate for the MC part of the exam (82.2%) and the pass rate for the MC part of the combined tests (71.2%) do not seem significantly different than for 2004-06 (77.9% and 71.6% respectively).

Table 7B: test MC & exam MC 2007-08

% of students by row	Pass exam MC	Fail exam MC	Row total
Pass test MC	92.6%	7.4%	100%
Fail test MC	56.5%	43.5%	100%

Of those who passed the test MC, 92.6% went on to pass the exam MC, a slight increase over 2004-06 (88.2%). However, those who failed the aggregate test MC still had roughly a 50/50 chance of passing the subsequent exam MC (a slightly higher 57/43). This pattern indicates that the test MC were a lot harder than the exam MC. Hence, passing the tests would almost guarantee a pass in the exam. But, a fail in the test MC wouldn't necessarily predict failure in the exam MC – in fact over half the people that failed the test MC managed to pass the exam MC.

2007-08 test short-answer vs. exam short-answer

Here we take only the SA part of the 2 tests (unweighted total) and the SA part of the exam; the marks for SA parts of both tests have been increased over the second period. The population includes only those students who completed both tests and the final exam ($n = 775$).

Fig.7.

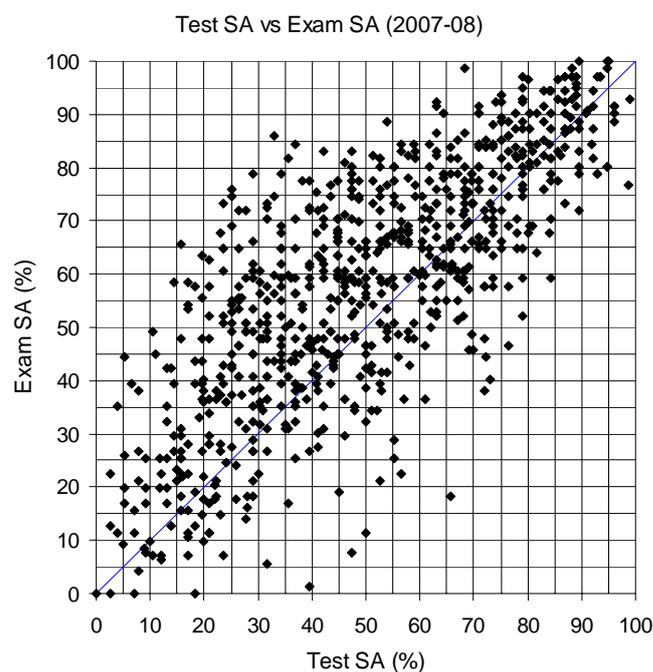


Table 8A: test SA & exam SA 2007-08

	mean mark	pass rate
test SA	50.3	50.2%
exam SA	58.3	65.6%
% students scoring more in exam SA than tests SA = 69.3%		

After the restructure of Econ 101, the two SA assessments are less equally spread above and below the $y = x$ line, with 69.3% of the population above. The result is the opposite of the second period: SA marks now are harder to gain in the tests than the exam, with means of 50.3 and 58.3 respectively. The pass rate for the SA part of the exam increased significantly over the previous period from 50.6% (2004-06) to 65.6%. We see three reasons why the pass rate for the SA part of the exam (65.6%) is higher than the pass rate for the SA part of the tests (50.2%). First, the exam allows students to choose 1 of 2 options to answer, while all test questions are compulsory. Second, about 2/3 of the exam SA cover topics students have already studied for in the 2 tests, and students can be expected to do better the second time around. Third, borderline students just failing the tests have an incentive to increase their effort for the exam.

Table 8B: test SA & exam SA 2007-08

% of students by row	pass exam SA	fail exam SA	row total
pass test SA	89.7%	10.3%	100%
fail test SA	41.5%	58.5%	100%

Increasing the quantity of short-answer marks in the tests appears to have significantly improved the robustness of a successful test result, as the exam SA pass rate for those who had passed the test SA went from 72.8% to 89.7%. (This is only just lower than the corresponding rate for the multiple-choice of 92.6%.) At the same time, the robustness of a fail test result fell. The exam SA pass rate for those who had failed the test SA rose from 25.2% to 41.5%; this is comparable to the corresponding rate for the multiple-choice, whereas for 2004-06 the exam SA pass rate for those who had failed the test SA was only one half the corresponding rate for the multiple-choice (25.2% vs. 51.9%).

Here, compared to the SA test/exam relation 2004-06, the top-left/bottom-right diagonal is less clear: the test SA passes result in exam SA passes at much higher rate than the test SA fails result in exam SA fails. This is consistent with the increase in the difficulty of the test SA and decrease in the difficulty of the exam.

Aggregate Tests 2007-08

“Aggregate tests” refer to the weighted coursework mark for both tests (combined MC and SA). The aggregated (weighted) test mark puts more weight on written answers for the SA questions and less on the MC questions by putting more weight on test 2 which has a larger SA component; in this period the weight for SA was increased over 2004-06. The population is all students who sat both tests and the final exam ($n = 775$).

Fig.8.

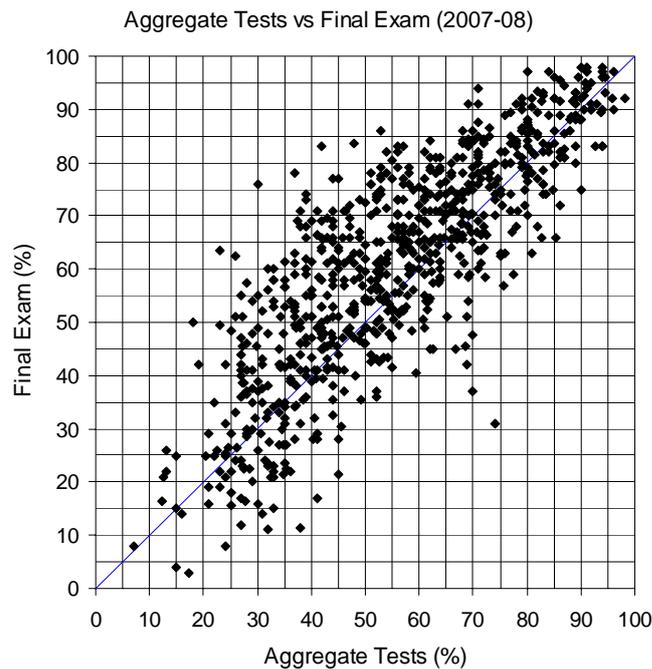


Table 9A: aggregate tests 2007-08

	mean mark	pass rate
aggregate tests	55.4	59.2%
exam	60.9	71.4%
% students scoring more in exam than aggregate tests = 67.3%		

The two assessments are fairly evenly spread above and below the $y = x$ line, but unlike 2004-06, the aggregate tests are now slightly harder than the final exam. With the increase in the proportion of test marks allocated to SA, the aggregate test means and pass rate fell from 56.7 and 67.0% (2004-06) to 55.4 and 59.2%. However, over the same period, the exam pass rate rose from 62.1% (2004-06) to 71.4%.

Table 9B: aggregate tests 2007-08

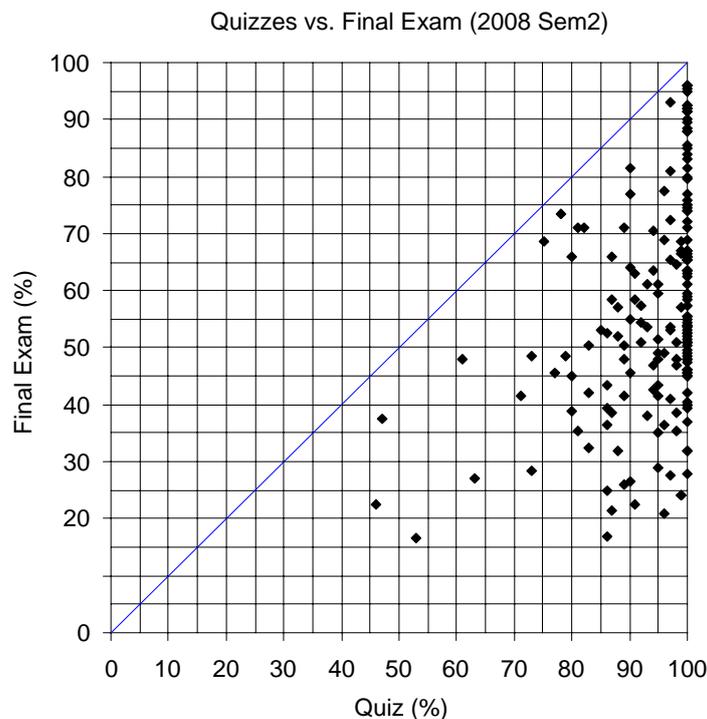
% of students by row	pass exam	fail exam	row total
pass tests	93.4%	6.6%	100%
fail tests	39.6%	60.4%	100%

As evident in the SA test/exam relationship, compared to the previous period, the combined effect of more difficult tests and an easing in the exam is an increase in the pass exam/pass test ratio (from 81.0% to 93.4%) and a decrease in the fail exam/fail test ratio (from 76.2% to 60.4%). Those who pass the harder tests were more likely – now almost certain – to pass the easier exam, and vice versa with now 39.6% of those failing the weighted tests actually passing the final exam.

On-line Quizzes (2008 Semester 2)

“Quiz” refers to the average of the best 10 quiz marks (of the 12 available quizzes), and “Exam” refers to the total mark for the final exam. (There was, as expected, a marked drop off in quiz attempts after the 10th quiz. A lot of students had gathered sufficient marks from the first 10 quizzes and didn’t see the need to score highly in the last two quizzes.) The population is all students (n = 187) over 1 semester who completed at least 10 of the 12 quizzes and sat the final exam.

Fig.9



As quizzes have unlimited attempts, all results are virtually identical to those of the take-home assignments.

Table 10A: quizzes and final exam 2008 semester 2

	mean mark	pass rate
quizzes	94.2	98.9%
exam	56.60	61.5%
% students scoring more in exam than quizzes = 0%		

Table 10B: quizzes and final exam 2008 semester 2

% of students by row	pass exam	fail exam	row total
pass quizzes	62.2%	37.8%	100%
fail quizzes	0%	100%	100%

Only 1.1% of students failed the quizzes after completing all the required quizzes (i.e. 10 or more). This is because people who complete all quizzes are more conscientious and make an effort to get a reasonable mark. Given the mark is also based on their best 10 quiz results it is even easier to get a pass mark.

As a means of formative assessment, quizzes are more efficient than assignments. Both quizzes and assignments are 'easy' marks and students can gather very high marks by putting in a reasonable effort. Because of this, neither assignments nor quizzes perform well as summative assessment (as shown by their poor prediction of exam passes) and hence the reduced weighting on quizzes lowers the bias caused by the 'easy' marks.

3.4. Participation and exam pass rates

Participation is recorded as the completion rate after the addition of aegrotats, i.e. including students who did not complete but would have if they were able.

As a necessary but not sufficient condition for passing, a student's participation in the assessment signals only that passing remains a possibility. But failure to participate in any particular assessment is a definitive signal of failure for that piece of the course. The number of formative assessments completed signals the student's level of participation. Completing less assignments or quizzes signals

less engagement which, without a dramatic change in behaviour, will almost certainly result in failing the course; completing more pieces signals more engagement, which increases the chances of passing the course. The difference participation in either assignments or quizzes makes to exam pass rates speaks for itself.

Fig.10

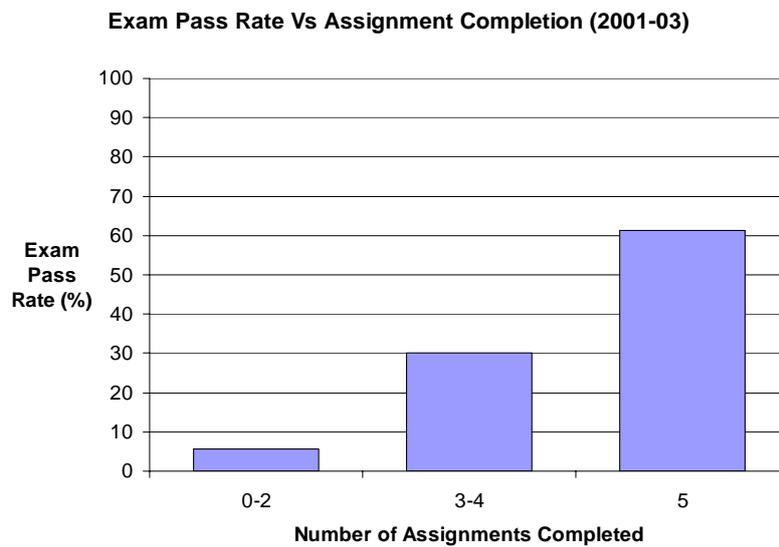
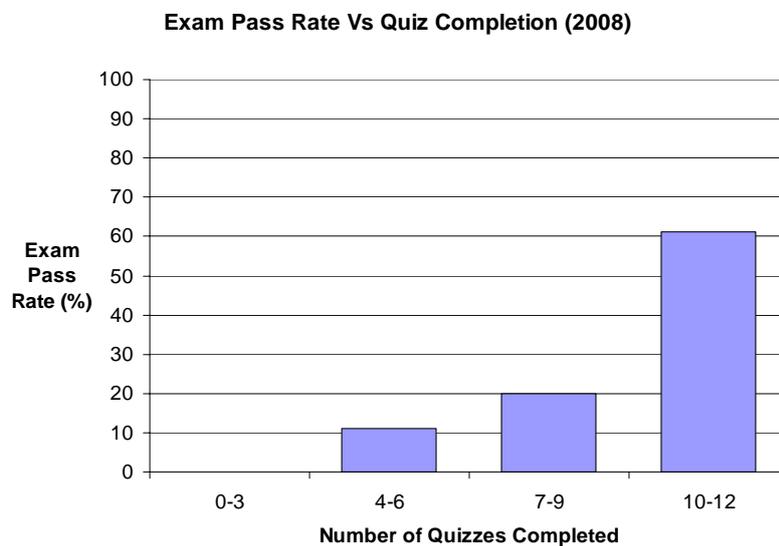
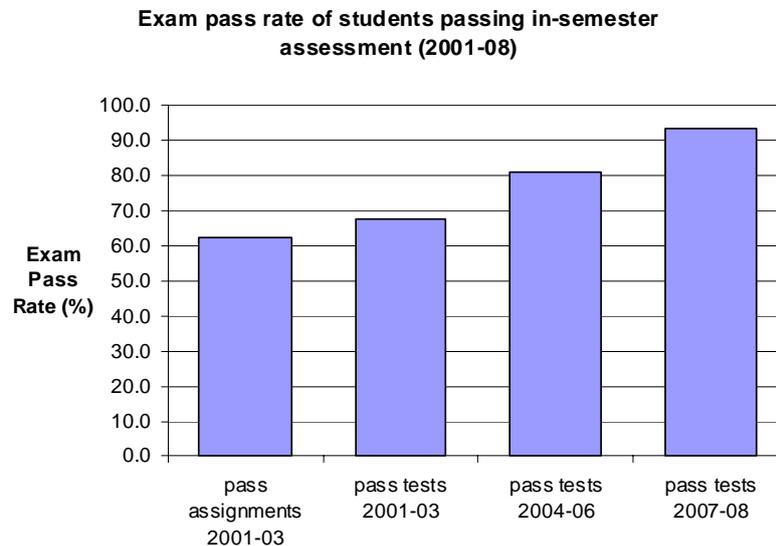


Fig.11



4. Conclusion

Fig.12



Over time we have been able to increase the exam pass rate of students passing in-semester assessment by making that assessment more like the exam in format and intensity. We have attempted to build a formative or learning experience into the summative assessment. Test 1 introduces SA questions and provides feedback on students' answers to allow lessons to be learned before test 2; test 2 increases the quantity of SA questions, gets closer to the exam in length and format, and provides further feedback and adjustment to their skills before the exam; then the exam is the most comprehensive – and final – assessment.

Assignments have traditionally been the favoured formative assessment but our experience indicated two things. First, that if students were handing in their own work (and there was some evidence of copying), many were learning little from the exercise. Second, that we were learning little about the students by marking assignments. Take-home assignments proved both insecure and unreliable as assessment. A further problem (of our own making) was that the excessive weighting given to assignments in the final mark created difficulties for the other forms of assessment. Most students started with nearly 30% after the assignments, and to pass needed only to get 20 out of the remaining 70 marks from the tests and exam. This compromised our ability to get a reasonable course pass rate and discriminate between average and excellent students. To try to reduce this problem, we introduced a minimum exam mark requirement for final grades; students could not pass the course without scoring at least 40 in the exam. This was intended to filter out students with high assignments marks who could not perform in the exam. However, the constraints on the grading spread remained. After 3 years we decided to simply drop the assignments and add short answer questions to the tests. We

conclude that 12 on-line quizzes worth 10% of the final grade are much more efficient and appropriate than 5 take-home assignments worth 30% of the final grade.

The questions asked in this paper are interrelated. The predictive power of in-semester assessment for the final exam goes 2 ways. The easier the first assessment is compared to the exam, the less meaningful a pass in the first is for a pass in the exam, and the more meaningful a fail in the first is for a fail in the exam. Conversely, the harder the first assessment is compared to the exam, the more meaningful a pass in the first is for a pass in the exam, and the less meaningful failure in the first is for a fail in the exam. So passing comprehensive tests is a good indicator of exam success, and failing easier formative assessment (assignments or quizzes) is a good indicator of exam failure, which is relevant for identifying where intervention may be needed.

Open-book assessment (assignments and on-line quizzes) and tests with only multiple-choice questions all proved limited as predictors of success in the subsequent examination, though more robust as predictors of examination failure. Tests that required written answers were better indicators of examination performance than tests with multiple-choice only, but the tests did not perform as well as assignments a means of formative assessment.

How well in-semester assessment provides incentives that keep students engaged in learning throughout the semester depends on whether it is summative or formative. Coursework that does not duplicate the scope and intensity of the final examination is not a substitute for a final examination, but is a complement which engages students throughout the semester. Awarding marks for coursework provides an incentive to study and reduces the crowding out of assessment by the demands of other courses.

There is a trade-off between formative and summative assessment. The most useful balance between formative and summative assessment provides formative assessment which creates incentives to study, and summative assessment which allows both students and teachers to gauge progress. Formative assessment is intended to engage students and help direct their learning. If this assessment is too difficult, students can be discouraged and drop out before the end of the semester. With formative assessment, the bar will necessarily be lowered if students are to be encouraged by success; hence passing formative assessment signals little more than a student's participation in the course, although failing formative assessment suggests disengagement and is a signal of likely failure in the later summative assessment.

We regard the most useful balance between summative in-semester assessment and the final exam is to make the tests harder than the exam, which is both better as an incentive for study and for fairness than

having the exam harder than the tests, which sends the wrong signals. Students passing the tests can reasonably expect that by continuing their current performance they should pass the exam; it is unfair to allow students the impression that they will pass if in fact they will not. Students would much prefer to fail difficult tests and, after applying themselves, find they pass the exam, than the reverse.

5. Appendix 1 – Assessment details

MC: multiple choice questions. All Multiple-choice questions have 5 options, with 1 mark for a correct and 0 for an incorrect answer; the expected mark with random choice is 20%.

SA: short-answer questions, including definitions, descriptions and explanations, drawing and interpreting diagrams, and calculations.

All Assignment, Test and quiz questions are compulsory. The Exam MC questions are compulsory, but the Exam SA questions allow students to choose 1 of 2 options.

Coursework is the weighted sum of in-semester assessment. Over all 3 periods the final mark consisted of 50% coursework plus 50% final exam.

2001 to 2003 (6 semesters)

Piece	Type	Marks	Time allowed	Weighting
5 Assignments	SA	18	2 weeks	$5 \times 6\% = 30\%$
2 Tests	MC	30	50 minutes	$2 \times 10\% = 20\%$
Final Exam	MC SA	29 71	190 minutes	50%

2004 – 2006 (4 semesters)

Piece	Type	Marks	Time allowed	Weighting
Test 1	MC SA	30 10	90 minutes	20%
Test 2	MC SA	30 20	120 minutes	30%
Final Exam	MC SA	29 71	190 minutes	50%

2007– 2008 (4 semesters)

2007 – 2008 Semester 1

Piece	Type	Marks	Time allowed	Weighting
Test 1	MC SA	26 14	90 minutes	20%
Test 2	MC SA	26 24	120 minutes	30%
Final Exam	MC SA	29 71	190 minutes	50%

2008 Semester 2

Piece	Type	Marks	Time allowed	Weighting
12 Quizzes	MC	10	no limit	10% average of best 10 quizzes
Test 1	MC SA	20 20	90 minutes	15%
Test 2	MC SA	20 30	120 minutes	25%
Final Exam	MC SA	29 71	190 minutes	50%

NOTE: We ignore the small increases in test short answer components and reduction in test course weighting in 2008 Semester 2 and include this with 2007 Semester 1 to 2008 Semester 1.

6. Appendix 2 – Tables

Tables C show the four possible pass/fail combinations by percentage of total students who completed all the relevant in-semester assessment and sat the final exam for the 3 periods.

Table 2C: total assignment marks and final exam marks 2001-03

% of all students	pass exam	fail exam	row total
pass assignments	61.0%	37.0%	98.0%
fail assignments	0.3%	1.7%	2.0%
column total	61.3%	38.7%	100%

Table 3C: MC only tests 2001-03

% of all students	pass exam	fail exam	row total
pass tests	51.9%	24.9%	76.8%
fail tests	3.5%	19.7%	23.2%
column total	55.4%	44.6%	100%

Table 4C: test MC & exam MC 2004-06

% of all students	pass exam MC	fail exam MC	row total
pass test MC	63.1%	8.5%	71.6%
fail test MC	14.8%	13.7%	28.5%
column total	77.9%	22.2%	100%

(Note not all totals tally due to rounding.)

Table 5C: test SA & exam SA 2004-06

% of all students	pass exam SA	fail exam SA	row total
pass test SA	38.8%	14.5%	53.3%
fail test SA	11.8%	35.0%	46.8%
column total	50.6%	49.5%	100%

Table 6C: aggregate tests 2004-06

% of all students	pass exam	fail exam	row total
pass tests	54.3%	12.7%	67.0%
fail tests	7.8%	25.1%	33.0%
column total	62.1%	37.9%	100%

Table 7C: test MC & exam MC 2007-08

% of all students	pass exam MC	fail exam MC	row total
pass test MC	65.9%	5.3%	71.2%
fail test MC	16.3%	12.5%	28.8%
column total	82.2%	17.8%	100%

Table 8C: test SA & exam SA 2007-08

% of all students	pass exam SA	fail exam SA	row total
pass test SA	45.0%	5.2%	50.2%
fail test SA	20.6%	29.2%	49.8%
column total	65.6%	34.4%	100%

Table 9C: aggregate tests 2007-08

% of all students	pass exam	fail exam	row total
pass tests	55.3%	3.9%	59.2%
fail tests	16.1%	24.7%	40.8%
column total	71.4%	28.6%	100%

Table 10C: quizzes and final exam 2008 semester 2

% of all students	pass exam	fail exam	row total
pass quizzes	61.5%	37.4%	98.9%
fail quizzes	0	1.1%	1.1%
column total	61.5%	38.5%	100%