# On-line delivery and marking of Excel based assessments

**David Whigham** and **John Houston**
d.whigham@gcal.ac.uk j.houston@gcal.ac.uk
Department of Economics
Glasgow Caledonian University

## Abstract

The workshop will demonstrate a series of Excel files and worksheets that assess a variety of Excel based questions suitable for a range of undergraduate teaching. Topics range from introductory generic Excel skills (sorting, filtering etc.) to intermediate mathematical modelling and statistical analysis. Although the nature of the assessments derives primarily from Business Studies and Economics, the principles being assessed will be germane to many other cognate areas.

The novel aspect is that students are presented with an Excel file consisting of a number of individual sheets containing the questions to be answered, and instructions to return their answers to specified cells. The marking sheet (which is hidden and password protected) then assesses these answers and awards the appropriate mark as defined by the marker. This gives respondents the opportunity to correct any answer if less than full marks have been awarded by the marking sheet. The marking sheet has been thoroughly de-bugged and live tested, and will not be affected by responses that produce error messages.

Additionally all of the prepared questions have been collected into a question bank, that easily allows instructors to select only those questions that they want to include in the assessment they are preparing. A wide variety of customised tests containing both different questions on the same topics, and/or different questions on different topics can therefore be created in minutes.

There is also an option to make each assessment file monitor students' response behaviour throughout the duration of the examination and thereby provide a rich data source on the extent to which the ability to see the mark awarded and, if necessary, correct the answer, influences the manner in which students approach the assessment.

A number of other features concerning security, assessment integrity and marking tolerances will be discussed and displayed at the workshop.

## Introduction

As members of the Department of Economics at Glasgow Caledonian University the authors have been responsible for the provision of a number of modules, at a variety of levels, that both teach and make use of Excel spreadsheet skills. Some of these simply provide instruction on how to use basic Excel facilities and routines such as formula writing, copying, charting, sorting and filtering in the context of a variety of generic business problems. Others use these and some more advanced routines in the areas of mathematical and statistical business modelling.

It soon became clear that with student numbers in excess of 300 per semester, and rising, the marking tasks were becoming extremely onerous. Consequently, it was decided to develop a series of Excel workbooks for each of these modules, which would be able to mark the students' assessments on-line.

At first, these automatic procedures were on a modest scale, but as experience was gained we were able to expand the range of questions that could be marked in this way into an assessment bank of considerable size and variety.

Of course, it was recognised that not each and every question in our assessments would be deemed appropriate for the module(s) that other instructors deliver. So, once a sufficiently large bank of questions was created, it was then decided to include them in a customised menu suite, so that examiners could easily pick and choose those questions that they wanted to include in their assessments. This was done using Visual Basic (VBA), and simply requires that the test creator select those questions from the bank that are **not** to be included in the assessment. Thereafter the program excludes them and reformats the file into the standard assessment format. A wide variety of customised tests containing both different questions on the same topics, and/or different questions on different topics can therefore be created in minutes.

Additionally, since some of the assessments were of a statistical nature, we also created an Excel file that could calculate various probabilities for all of the commonly used probability distributions. This file could be kept in background while students were attempting the assessment and then accessed when a question required the calculation of a probability. It therefore constituted what we decided to call an “on-line Murdoch and Barnes”.

## The tests

At the start of each semester files pertaining to each module are mounted on the intranet and numerous practice assessment files and other resources made available. Students are shown how to access these and copy files to their own disc. These practice assessments are also automatically marked, and therefore give students the opportunity both to practise answering

questions and also to familiarise themselves with the protocols that will be required in the “real” assessments.
The opening sheet for a typical assessment is shown in **Figure 1**.

**Figure 1**

| | | | | |
|---|---|---|---|---|
| FIRST NAME | | SURNAME | | |
| THE COMPUTER WILL MARK THIS EXAMINATION. | | | | |
| CONSEQUENTLY, YOU **MUST** LOCATE YOUR ANSWERS TO THE QUESTIONS IN THE CELL | | | | |
| OR CELLS DEFINED IN THE INSTRUCTIONS. | | | | |
| IF YOUR ANSWER IS LOCATED IN THE WRONG CELL THEN IT WILL BE REGARDED BY | | | | |
| THE MARKING PROGRAM AS WRONG - EVEN IF IT IS ACTUALLY CORRECT | | | | |
| | | | | |
| QUESTIONS AND INFORMATION ARE IN PINK COLOURED CELLS: | | | | |
| YOUR ANSWERS SHOULD BE SENT TO YELLOW COLOURED CELLS | | | | |
| SOME OF THE SHEETS ARE **PROTECTED** TO PREVENT ACCIDENTAL | | | | |
| ERASURE. WHERE THIS IS THE CASE WORKING AREAS HAVE BEEN | | | | |
| CREATED IN BLUE COLOURED CELLS. | | | | |
| ANY WORKING THAT YOU REQUIRE TO DO **CAN ONLY BE** | | | | |
| **ENTERED TO THESE AREAS**. | | | | |
| | | | | |
| THERE ARE A TOTAL OF 14 QUESTIONS AND 110 MARKS IN THIS EXAMINATION PAPER. | | | | |
| **YOU MAY ATTEMPT AS MANY QUESTIONS AS YOU WISH**. | | | | |
| FULL MARKS (100%) WILL BE GAINED BY A TOTAL OF 80 MARKS. | | | | |
| ALL MARKS TOTALLING 80 OR LESS WILL BE EXPRESSED AS A PERCENTAGE OF 80. | | | | |
| | | | | |
| THE COMPUTER IS MARKING YOUR ANSWERS AS YOU COMPLETE THE QUESTIONS. | | | | |
| **IF YOU WANT**, YOU CAN HAVE YOUR SCORE TO DATE SHOWN IN THE SHEET CALLED | | | | |
| MARKS. | | | | |
| TO DO THIS YOU MUST ENTER THE LETTER Y TO THE GREEN CELL BELOW. | | | | |
| | | | | |
| YOU CAN TURN THE SCORE OFF BY ENTERING ANY CHARACTER OTHER THAN Y TO THE | | | | |
| GREEN CELL | | | | |
| | | | | |
| **NB YOU MUST SAVE THE FILE AFTER YOU HAVE MADE YOUR FINAL ANSWER TO EACH QUESTION**. | | | | |
| IF YOU CORRUPT THE QUESTION SHEET THEN CLOSE THE FILE **WITHOUT SAVING** AND | | | | |
| THEN RE-LOAD IT. YOUR CORRECT ANSWERS WILL BE INTACT BUT THE | | | | |
| DAMAGE THAT YOU DID WILL BE GONE. | | | | |
| THERE IS ALSO A HELP SHEET. CLICK ON IT TO OBTAIN INFORMATION ABOUT | | | | |
| WHAT YOU MIGHT HAVE DONE WRONG. | | | | |
| GOOD LUCK | **END OF INFORMATION** | | | |

It will be noticed that the ability to see the marks awarded is at the student’s discretion, with the default being that the marks will not be shown. This feature was incorporated after protracted discussion amongst the teaching team. Some felt that if a student were performing poorly then seeing that this was the case would act as a disincentive. On the other hand, seeing that the answer was wrong allowed the opportunity to correct and perhaps improve performance. Consequently, it was decided to give students the choice of seeing their marks, as well as the ability to change their minds at any time.

In a related vein, the authors also thought that seeing when a pass mark had been obtained might act as an incentive to under-achieve, while seeing that say 98% had been obtained could act as a spur to obtain the perfect score. These incentive and disincentive effects were of interest to the authors, and we report some tentative findings later in the paper.
A typical question from an introductory Excel skills module is shown in **Figure 2**.

**Figure 2**

| SORT THE FOLLOWING LIST OF SHARE DATA IN ASCENDING ORDER OF SHARE PRICE. | | | | | | | |
|---|---|---|---|---|---|---|---|
| SEPARATE ANY TIED RANKS FIRST OF ALL IN ASCENDING ORDER OF DIVIDEND AND | | | | | | | |
| THEN IN DESCENDING ALPHABETICAL ORDER. | | | | | | | |
| | | | | | | | |
| Share name | Share price (£) | Dividend (£) | | | | | |
| Spoolers | 0.95 | 0.19 | | | | | |
| Regal Bank | 1.56 | 0.4 | | THERE IS A BACK UP DATA SET | | | |
| PR&W | 2.03 | 0.49 | | IN A26:C41 IN CASE YOU MAKE | | | |
| Pandora | 0.95 | 0.19 | | AN UNFIXABLE MISTAKE. | | | |
| Pan-europe | 2.7 | 0.4 | | | | | |
| MBI | 0.99 | 0.18 | | | | | |
| IFI | 0.95 | 0.19 | | | | | |
| Gloxo | 3.21 | 0.5 | | | | | |
| Devoors | 1.5 | 0.31 | | | | | |
| Credit Alpha | 12.21 | 1.65 | | | | | |
| Costcut | 2.7 | 0.35 | | | | | |
| Commercial Plus | 1.67 | 0.21 | | | | | |
| Cathedral Life | 2.43 | 0.34 | | | | | |
| Brutus | 2.7 | 0.35 | | | | | |
| AM&M | 1.5 | 0.31 | | | | | |
| Agricola | 0.65 | 0.12 | | | | | |

## Security and robustness

The marking sheet is contained within the actual assessment file, but is hidden, password protected and formatted to be invisible to the student. Without doubt an Excel expert could reveal this sheet and, with it the correct answers. However, since the assessments are not targeted at Excel experts, and since the actual assessable tests are invigilated this was not deemed to be a serious problem.
Nevertheless, a number of further difficulties had to be resolved before we felt confident that the marking sheet was secure. The following are just a few :-

- Student answers that produced Excel error messages had to be trapped and ignored by the marking sheet so that an eventual score for the correct answers could be computed (a single error message in a list of numbers prevents totalling).
- The default blank response to a question had to be differentiated from zero so that a correct answer of zero would not be marked as correct with the default response.

Both these and other difficulties were satisfactorily resolved by judicious use of formulae in the marking sheet.
However, the sensitivity of the tests to either intentional or unintentional corruption by the student was a more worrying problem.
At base, the issue of robustness stems from the inability to protect many of the sheets when they require the use of Excel routines such as Sorting, Filtering, Solving or Data Analysis.

In the first stages of the project, despite intensive testing by the authors, it was still found that both the question and marking sheets could become corrupted in a variety of ways when subjected to student use. Fortunately however, we were able to monitor students who were doing the practice assessments and thereby discover a wealth of potentially destructive processes that could then be incorporated into the safety checks prior to the 'real' assessments.

This was done in a variety of ways that usually involved using VBA to remove those menu options and Toolbar buttons that could otherwise cause damage, but were not essential to the assessment. Then, so that a student who was VBA literate could not put them back on again, the VBA was removed. As can be imagined, this added considerably to the time involved in composing the assessments, since after removing the visual basic toolbar we would frequently find the need to use VBA again - a classic case of painting oneself into a corner.

However, even with this fairly high level of protection established it was decided to include a restore button in each question sheet that would clone the original question in the event of mishap. Inevitably this meant that the file size increased, but even for an assessment containing 20 questions this never exceeded 200Kb.

Finally, since all of the 'real' assessments would be invigilated by experienced Excel tutors, we felt confident that any untrapped problems that arose could be rectified with the minimum amount of disruption to the student. Extra time would, of course, be allowed where this happened.

## Limitations

It will be clear that the marking sheet can only assess answers to questions that require numeric, non-binary responses, and this raises two issues.
First, how 'right' must the answer be before it is awarded full marks?
Second, how can the marking sheet assess responses that are non-numeric – charts or comments for example?
In regard to the first issue, our original approach was to employ zero tolerance of arithmetic deviation from the correct answer. This created difficulties however.

The default decimal place and type formatting frequently varied across the computers being used for the assessment, so that the marking program could return a zero mark for an answer that was only fractionally different from the

correct answer, or expressed in a different format. This was particularly true in problems involving use of the Excel Solver.
The fix we devised was to include a hidden auto open front end in the question files that ensured homogeneous Excel configuration when the assessment commenced.

Nevertheless, we still had to address the question of whether, given homogeneous configuration, an answer of say, 1.2 to a question for which the correct answer was say, 1.25 was close enough to warrant any marks.
Our eventual decision reflected our backgrounds as economists. We concluded that sometimes it was and sometimes it was not. To be more precise, we thought that where the question called for an extensive series of calculations that could each feed their error potential into subsequent calculations, then a margin of 5% error around the correct answer could be allowed. (But see our comments under Customisation as to how this could be changed). Thereafter, an answer that was outside 5% error, but within 10% would be awarded half marks. Any answer outwith a 10% error would receive no marks.
The logic of this approach is clearly that to come so close to the correct answer implied that the method(s) employed must have been more or less correct.
On the other hand, there were question types where even the slightest mistake would cause severe problems in real life - sorting being a classic example.

Regardless of the number of times that the tutor stresses the point, numerous students will not include the **entire** database when they are required to sort upon a particular field. They sort on that field **only** and therefore attach sorted data from one field to unsorted data for the remainder of the fields, thereby destroying the integrity of the entire database. Cases such as these were viewed as deserving zero marks.

As will be seen from the actual question sheets, there were a number of questions that required students to enter the **text** version of the formula that they used (without any **redundant** dollar signs). The latter requirement ensured that the marking program could assess these responses without recourse to a variety of dollar permutations to check all of the possible entries that would actually work.

The reason for requiring the text version was to ensure that students could write appropriate formulae and then copy them, rather than simply use a calculator on each and every required calculation. With regard to the issue of non-numeric responses we could do nothing within Excel that would assess any comments that were required. These had to be manually assessed by the tutor.

However, in the crucial area of charting we devised a number of questions that required, for a specified problem, that the student enter their responses to each stage of a simulated Chart Wizard dialogue. If all of their responses were correct (as assessed by the marking sheet) then the eventual chart

*should* have been correct. This worked very well, but was still unsatisfactory in the sense that the eventual chart could have been created in a variety of other ways that were semi-acceptable. We could do no more than accept a fully correct answer as marked by the system, or ask the candidate to show the actual graph and then mark it manually.

In any case, because of the nature of the software in our main teaching labs we had no alternative but to go round each student and record their final mark manually at the end of the assessment. At this stage we would then give credit for anything that they had done but which had been zero marked by the marking sheet. This was not unacceptable since as we have already said, the marking sheet is not able to identify credit-deserving attempts that are outside the tolerance limits, and so to a large extent the manual marking only applied to questions that had been zero marked.

## Efficiency

The efficiency of the system is easy to compute.
With an average of 300 students taking an average of 16 questions in formal assessments this would create 4800 questions to be marked.
Prior to the introduction of this system, manual marking took approximately three minutes per question, giving 24 staff marking hours.

With the on-line marking we can assess 300 students on five modules (including manual marking of graphs and required comments) with two staff, in 30 minutes per module. This gives five staff marking hours, and represents efficiency saving of almost 75%.

There are, of course, the staff hours involved in creating these assessments, but we now regard these as a fixed cost, and they are zero cost rated for those assessors who decide to use the material that we have produced.
This high level of efficiency is, of course, dependent upon current student numbers and lab availability. At the moment we are able to access six labs that each have twenty computers, and serendipity has decreed that our largest module comprise 110 students. We are therefore able to invigilate the entire module with four tutors (since two pairs of labs are interconnected). On average this two-hour assessment is completely marked in less than three staff hours.

## Customisation

The marks awarded to each question clearly reflect our perception of the difficulty of the question in relation to the nature and extent of tuition received by our students. Consequently, we recognise that users of the system may well want to alter these parameters. This is easily done from the marking sheet and will automatically spread into the question sheets, the marks sheet and also the front-page where the marks details are displayed. (Where it was impossible to protect a question sheet because an Excel routine had to be used, it is clearly feasible for the student to change the marks awarded to each question or question part. However, this is a virtual change and will not

affect either the mark sheet or the marking sheet, which are both fully protected).

It is also an easy matter for the instructor to change the 'threshold' mark out of which the assignment is to be marked (80 marks in Figure 1) and this will then spread into the calculation of the percentage mark awarded.
Wherever possible the correct answers have been calculated in the marking sheet as 'live' formulae. This allows the assessor to alter the data in the question and be secure in the knowledge that the marking sheet will have adjusted the correct answer automatically.

This said however, there are clearly question types where the correct answer must be fixed data items, since no formulae are employed. Sorting and Filtering are the most obvious cases.

Here, data variation by the assessor will require that the correct answers are computed and then pasted into the appropriate range of the marking sheet. As regards tolerances on the correctness of the answer, then these too are at the assessor's control from the marking sheet, and clearly need not be the same for each question, or question part.

## Experience to date

To date, once the teething troubles were eliminated, student response has been excellent.
We delivered a questionnaire to students that focused on the merits or demerits of being assessed in this way.
Their responses were almost unanimously in favour of this form of assessment and of the ability to see their marks. This was further confirmed in the actual assessments by the fact that very few students elected **not** to see their marks (and when they did this is was almost always due to oversight). Students were also highly appreciative of the practice tests, with one of the commonest comments being that they could use them at home and receive immediate feedback on their performance.

These practice tests contained both on-line marking of attempts to solve the questions as well as fully worked and explained solutions that were only observable once the student had made an attempt at a particular question.

## Incentive and disincentive effects

As mentioned previously, the teaching team was interested to know how the ability to see their marks influenced student behaviour in the assessments. On the basis of both visual evidence and a rudimentary automatic monitoring device, we were able to make some tentative observations.

There was considerable evidence of under-achievement as measured by the number of students who left the assessment early once the pass mark had been achieved. Our estimate is that approximately 30% fell into this category,

although it is impossible to say whether some of these students simply could do no more of the questions and therefore left without having under achieved. There was strong evidence that the ability to see their mark encouraged correction, and that more often than not, this correction was successful.

However, there were a worryingly large number of students who ‘stuck’ at a particular question and would not leave it because they had not received full marks. The tutors were all instructed to remind such students of the Principle of Diminishing Returns, but even then, there was evidence that some students stubbornly declined this advice.

On average approximately 7% of students left their assessment prior to its scheduled end without having achieved the pass mark. Of these however, over 85% had made no attempt at more than two thirds of the questions, and so could not have been put off by the ability to see their marks.
Encouragingly, there were at least 20% of the cohort, who managed to turn a failing mark into a pass in the last 10% of the examination time, simply by retrying questions that they had got wrong and then checking the marks awarded.

Finally, roughly 8% of students displayed what can only be called doggedness to obtain the perfect score. The number of students who had obtained 90% or more within the first three-quarters of the assessment period, and yet to the end remained measured this.

We realise that these findings are to a large extent anecdotal, but have decided to devise a more rigorous statistical approach as part of our future developments.

## Future developments

We recognise that there will always be a need for more questions to reflect variations in syllabus content and Excel capabilities, as well as the requirements of new modules. What is more time consuming however, is the need to incorporate these new questions into the menu suite package.

Also, given the decision to include clones of each question sheet in each assessment file, the size of menu suite file expands dramatically.
The obvious solution is to have a set of menu suites from which assessors can choose their preferred questions. As yet however, we have been unable to devise a satisfactory method of finally consolidating choices from a series of separate files into one single assessment file. But we are confident of being able to do this soon.

As was said earlier, at the moment we are content to record the students’ final marks manually, and see little possibility of ever being able to withdraw some form of human contact at the end of the assessment. This said however, it would clearly be desirable to be able to have the tutor check the final mark, add any manual marks and then have the final mark recorded automatically in

a marks file for the entire module. (At the moment we have to transfer the final recorded mark manually into our module marks file).

Both the authors and Excel has the capability to do this, but at the moment the teaching labs do not have a sufficient number of Windows NT machines to allow a module marks file to access each individual student's assessment file. However, once all of the labs are converted to NT, then this will be our next development.

However, a much more ambitious extension relates to the process of monitoring student behaviour during the assessment. At the moment, we can include in any assessment file a sheet that can track, amongst other things, the time spent on each question, the time taken making corrections to questions that have not obtained full marks, and movement patterns between questions.

Nevertheless, there are a number of extensions to this monitoring file that require further development before it can supply all of the information that we require to perform a rigorous statistical analysis of incentive and disincentive effects.

## Conclusion

We hope that this paper has provided a cogent explanation of the processes employed in developing what we believe to be a novel method of marking Excel based assessments.

All of the question files, as well as the customising program are freely available by contacting the authors.