Appendix 2

Curriculum review methodology

Overview
The curriculum review analysed 174 economics modules at seven Russell Group universities: Cambridge University, University of Sheffield, Queen’s University Belfast, University of Manchester, London School of Economics, University of Glasgow and University of Exeter. It drew on two sources – module course outlines and past exam papers – in an attempt to identify empirically the content and style of undergraduate economics education in the United Kingdom.

The bulk of the research for the curriculum review was carried out in the 2014/15 academic year. We approached 16 universities to ask if they would share their course outlines and past papers with us. Most said no, with only Queen’s University Belfast, Oxford and Cardiff saying yes (the last two were not included in the review since we received them too far into the process). This meant that most of the materials for the curriculum review were gathered by supportive students who were studying economics at the universities analysed and the whole process was significantly more time consuming than we had expected. As a result, our sample consists of universities for which we could obtain data, rather than a random sample or representative sample. If, as we argue in this book, the state of academic economics is of public interest then there should be more openness and transparency around what economics students are taught in publicly funded universities in the UK today, and universities should make many of these materials publicly available.

All course outlines in our sample were for the 2014/15 academic year. Exams in our sample were predominantly made up of final examinations for the 2013/14 academic year or sample papers provided to students. In very few instances papers from earlier years were used, as far back as the 2008/09 academic year. However, such papers had always been provided to students for exam practice, suggesting that they were still broadly similar to the final examinations that students would take. For logistical reasons, we generally could not gather mid-terms and
coursework in a systematic way. In the few instances when mid-terms could be obtained but final examinations could not, the mid-terms were used as a proxy for the final exams.

Course outlines

While the nature and comprehensiveness of course outlines varied considerably between universities, we used them to collate our data on which textbooks were used, which theories and models were taught and whether non-neoclassical perspectives were mentioned. We also used the course outlines to identify how much of each module’s assessment was made up of the final exam so that we could identify how much we were missing by not analysing mid-terms and coursework.

Exam past papers

Of the 174 modules we studied we were able to analyse past papers for 156. This is because some courses are 100 per cent coursework-based, but also because others could not be obtained. There is no reason to believe that the exams we could not collect would have had a particular effect on our results. While an academic from outside economics might assume that if coursework and mid-term exams were added to our results they would improve significantly, our analysis of course outlines and anecdotal evidence from economics students across the UK suggests that these parts of assessment are dominated by multiple choice and assessed problem sets that leave much to be desired. Furthermore, on average at the universities studied, final exams constitute 81 per cent of all assessment and so we are analysing the vast bulk of assessment.

We devised a methodology to understand what knowledge and skills economics education requires students to have in order to graduate and become a qualified economist. We are fully aware that, as with almost any research exercise, judgement is involved in selecting what to measure and carrying out the analysis, and we do not try to hide that. No doubt somebody else could design a curriculum review that arrived at different numbers (though we doubt anyone would conclude that economics education is highly critical and empirical). Our aim here is to state transparently the judgements and assumptions we made, demonstrate why they are reasonable and point out the limitations of our research.

We developed four categories so that we could allocate each exam question into a category and come up with a total for each module,
particular groups of modules like macro and micro, universities and the whole sample. These four categories are:

1) **Operate a model**: Operate a model questions ask students to focus on a particular economic model and to perform some task with it. It entails no evaluation of whether the model is appropriate nor its strengths or weaknesses. Such questions can be graphical, mathematical (game theory, constrained optimisation) or even, in some instances, descriptive. Different kind of operate a model questions include:

   a) State and explain a model’s assumptions (no indication that this should be done critically, otherwise marked as evaluative).
   b) Derive and describe a model – draw graphs, solve equations, explain the theory, e.g. ‘In the Diamond-Dybvig model, explain whether the same outcomes for consumption can be achieved by replacing banks with financial markets in which investment projects can be traded.’
   c) Manipulate a model: once derived, show how something works within the model, e.g. ‘Suppose there is VAT i.e. $t>0$. Show that the results you have derived in part (d) might not hold here. Explain your answer.’
   d) Operate a model to explain a real-world event – a question that asks students to relate the model to the real world, e.g. ‘How can the IS/LM model explain the 2008 financial crisis?’
   e) Mathematical computation: questions that were purely mathematical were also classified as operate a model. Sometimes mathematical questions are not related directly to a model; however, the key criterion in operate a model questions is that the student simply works through maths without evaluating it, which also applies to mathematical questions.

2) **Describe questions**: Describe questions simply ask students to write down something from memory. Crucially, describe questions require no form of independent judgement. Describe questions can ask students to describe:

   a) A policy, institution or event, e.g. ‘What happened to stock prices during December 2008?’, ‘What is a fully funded social security system?’
   b) An argument, e.g. ‘Describe Larry Summer’s theory of secular stagnation’, ‘Give three reasons why a higher labour income tax may lead to a lower level of GDP growth?’
c) A theory, e.g. ‘What is the Friedman rule?’, ‘Explain the Marshall Lemer Conditions’, ‘Explain the characteristics of an optimal currency area.’

d) A methodology: common in economic policy analysis module questions, these simply ask students to describe the processes needed to get a particular finding, e.g. ‘How did David Card’s paper “The impact of the Mariel Boatlift on the Miami Labour Market” go about showing the effects of immigration on wages?’

3) Evaluative questions: Evaluation requires some form of independent judgement. For a student to make an independent judgement there must be more than one possible answer or at the very least not a strictly defined ‘right’ answer. Evaluation questions can include:

a) Evaluate a policy or event, e.g. ‘What were the key causes of the 2008 financial crisis?’

b) Evaluate a theory or model, e.g. ‘How useful is the Solow Growth Model in understanding long-term economic growth in the United Kingdom?’

c) Evaluate an argument, e.g. ‘For the two and a half decades since 1980, unemployment has been systematically higher in Europe than the United States because Europe has continued to have stronger employment protection and more generous unemployment insurance than the United States. Discuss.’

d) Evaluate one theory or model against another, e.g. ‘Which better explains the nature of the 2008 financial crisis, Real Business Cycle Theory or New Keynesian models?’

e) Evaluate methodology, e.g. ‘How useful do you think regression discontinuity designs have been in investigating the effect of education on wages?’

4) Multiple choice: This kind of question is self-explanatory and easy to identify.

Weighting

Each exam question was assessed in turn to see which of our categories it best fitted into. If exam questions seemed to fit multiple categories, the marks were split evenly between these categories. An example of a half-description, half-evaluation question from the review is ‘Describe how a researcher would carry out the contingent
Curriculum review methodology

valuation method in the context of a proposed woodland restoration project. Provide examples of the types of valuation questions used and discuss their advantages/disadvantages in your answer.'

If the questions were clearly designated as having more marks for one part than the other then it was split ¾ to ¼ e.g. ‘Derive the Solow growth model and briefly comment on its applicability.’

Aggregating and reporting results

For individual exams, the percentage of each category was calculated by seeing what proportion of total marks within the exam paper were assigned to that category. For example, if a paper had section A worth 50 per cent of the paper with one 50-mark question, and section B was worth 50 per cent and had two questions that students had to choose between, each worth 50 marks, then the total marks for that paper would have been out of 150. If question A was fully ‘operate a model’ and both question Bs were fully ‘evaluative’, then the score for that exam would be 1/3 operate a model, 2/3 evaluation.

To get aggregated scores for each university, the results for each individual module were multiplied by a weight relating to the credits (or percentage of the year) that they were worth. To aggregate results for all universities combined, each university’s scores for each category were summed and then divided by the number of universities.

In Chapter 2 we presented the results for all exams and core micro and macro modules. Core macro and micro exams were classified as microeconomic and macroeconomic exams that were compulsory for the universities’ single honours economics course (BSc, or if this was not available BA). So for example, Advanced Macro at Glasgow in Year 3 was not included because not every student has to take it. In most cases, such categorisation was self-explanatory; courses are typically called Microeconomic I, Macroeconomics II etc. In some instances, discretion was used to determine that a module was actually a core macro/micro exam. For example, at Sheffield, in the first year students take the modules Economic Policy 1 and 2, the content of which is clearly the same as macroeconomic and microeconomic courses at other universities. In the instance of QUB, where modules focus on more specific topics (e.g. Economic Growth Theory, Game Theory with Economic Applications) it was decided there was not a clear core/macro syllabus and so they were not included in these figures.
Challenges

The major challenge of the curriculum review lay in the fact that without the mark schemes we (like the students sitting the exams) had to go with what the question asked and this meant that the boundaries between different categories were not always clear. However, in some cases fuzziness was more problematic than others. For example, the boundaries between ‘describe a theory’ and ‘operate a model’ can overlap, but this is less relevant to our overall findings because both test technical knowledge and the skills of memorisation and regurgitation of that knowledge rather than independent or critical judgement.

More importantly, it was sometimes hard to tell whether a question was asking the student to evaluate something by making an independent judgement or they were just being asked to reel off something from the lecture notes/slides. Likewise, it is difficult to tell whether the student is actually able to argue a point and come to an independent judgement or whether the question has a single right answer that precludes the possibility of independent judgement. For example, in the exam we analysed for the third year course International Trade at Sheffield (where a mark scheme was included), Question 1C read ‘Use your knowledge about trade policy to evaluate the following statement: “tariffs have a more negative effect on welfare in large countries than in small countries.”’ This question seems a perfect example of critical evaluation, but in the mark scheme it actually says that the answer is categorically ‘False’, with a set answer required that simply used standard neoclassical models.

Our response to such ambiguity was to give the benefit of the doubt if a question seemed as if it could possibly be evaluative. This includes questions that form the final sub-part of an ‘operate a model’-style question and raise queries such as ‘name one drawback of the model’ for very few marks. From the experience of the authors, our peers and students throughout the movement it is likely that these questions are asking students to remember a one-line answer from the lecture notes and that the marks given do not allow real independent judgement. However, we have defined them as evaluative regardless.

We gathered all the data for the percentage of ‘operate a model’ questions that made a link to the real world and found that the results were very low, supporting our claim that economics education is abstract and disconnected. Again there was a challenge in defining
what constitutes the ‘real world’. Does an operate a model question relate to the real world if it simply mentions income, tax or the NHS? What about if it uses real-world data or reports real-world empirical trends? Or does it only count if the student is required to demonstrate substantive knowledge of the real world (and if that was the case how would we know that was a requirement)? Our criterion was that the question had to indicate that it required some kind of real-world knowledge in order to be answered. On this criterion a model that uses data from the real world doesn’t count because it doesn’t require the student to know anything, while using a model to explain or evaluate some real-world event or trend does count because the student has to bring in their own knowledge to answer the question.

Criticisms of this review could focus either on how the categories have been defined (e.g. that the evaluate criterion is too demanding or weak) or the challenges of applying them consistently and rigorously. We fully acknowledge that there are reasonable arguments for defining categories differently that would give different results. The more interesting question regards the issue of application. Throughout this process we have been aware of the role of judgement in the review and have taken steps to ensure that our process is rigorous while recognising that it isn’t objective. In our experience, the process was not so precise that different people would find exactly the same result every time. However, the curriculum review was always conducted by one member of the team and then again by another, to make sure that results were within 5 per cent. We therefore want to be clear that the process was precise enough to support the inferences we have drawn.

We would welcome further research on economics education from any interested parties and would be happy to discuss the pros and cons of our methodology further.

**Econometrics review**

The econometrics curriculum review was carried out based on different criteria to the other economics modules, simply because the original criteria did not fit. It included econometrics modules from the original sample of seven universities plus two econometrics modules from Cardiff University which were available because the econometrics review was done at a later date. Exam questions were split into three categories:

1. Independent approach to a problem. Questions that give students an issue to investigate or some raw data and then ask them how
they would analyse it and why. E.g. ‘You have data on tuition fees and the number of UCAS university applicants from 1990–2010. You wish to understand how the rise in tuition fees has affected the number of applicants. Which method do you use and why?’

2. Abstract/theoretical econometrics. Questions where no actual data or problem is mentioned, but a student is either asked mathematical questions about statistical/probability theory, or asked to manipulate regression models and test statistics at a purely abstract level.

3. Interpreting a given result or applying a suggested approach. Questions where a student is given a particular set of results or econometric method and asked to apply, interpret or evaluate it. Examples would be performing hypothesis tests from a given regression, or discussing omitted variable bias (either in general or for a specific case).

Furthermore, where possible we recorded the percentage of the marks for the entire module, as opposed to just the final exam (given by the course outline), that were given for an independent statistical report or project done outside the exam room.

Finally, a series of yes/no questions were asked about each module, with both course outlines and exams used to come to a conclusion. These included ‘is the course mostly linear regression (more than two-thirds)?’ and ‘is there critical discussion of data, data sources, etc.?’