Thoughts on the OTS December 2012 document on a complexity index

Richard Baron, Institute of Directors, richard.baron@iod.com

1. There are some questions to ask at the outset.

1.a. What is to have its complexity measured? I take it the answer is legislation, and these thoughts reflect that assumption, but it is possible that complexity of the whole system – legislation, administrative procedures, the interface between the tax system and the commercial world as it would exist independently of tax, and so on – is meant.

1.b. What would one do with a measure of complexity? Taxpayers and their agents must face whatever the law happens to be, are not in a position to make it less complex, and are often not in a position to avoid transactions that would make complex legislation apply (nor should they be deterred from transactions by the thought that the tax system will cause hassle). The main use is therefore likely to be to give guidance as to areas that are priorities for simplification, and to get an idea of whether any proposed changes would yield substantial reductions in complexity.

2. Given that, the usability model does look like a sensible starting point. But I suspect that the first stage in using it is to think hard about the goals. It is worth having a specific goal in mind, and thinking about it in some detail, so as to give practical substance to the model.

2.a. The obvious goal is to find out the amounts and times of any transactions with HMRC that will be needed. Taxpayers also need to know the administrative rules – which forms to use, when to file online, the (physical or electronic) addresses to which forms and money should be sent – but such details tend not to be obtained from legislation, even though it is primary or secondary legislation that empowers HMRC to make the administrative rules.

2.b. Another goal would be to discover in advance the relative tax merits of achieving results in different ways – buying or leasing equipment, using a company or an LLP [limited liability partnership], etc.

2.c. A third goal would be to learn in advance what systems and records would be needed, in order to allow compliance with the legislation in due course. For example, what records of purchase price, purchase date and enhancement expenditure will be needed for CGT purposes?

2.d. These different goals might lead to different assessments of the complexity of a given piece of legislation. For example, goal 2.a. assumes that you already know what happened, with monetary values, and your only interest is in what did happen, not in what alternatives might have happened. Achieving goal 2.b., on the other hand, would require you to consider the general scheme of the legislation and what

the results would be under the various alternatives, probably without knowledge of the monetary amounts. Indeed, to achieve goal 2.b., you might have to consider several separate pieces of legislation.

2.e. Despite the scope for different assessments of complexity, it is perfectly plausible that legislation that rated as high-complexity by reference to one goal would rate as high-complexity by reference to the other goals, and likewise for low complexity. So the ordering of different pieces of legislation by complexity might be pretty much the same, whichever goal one kept in mind.

2.f. Even if this is so, I would suggest picking on one goal, and keeping that in mind throughout. 2.a. is the obvious goal to pick, especially since everyone has to work out the tax consequences of their transactions, while some people do not need to consider alternative transactions, and achieving goal 2.a. requires keeping the records mentioned in 2.c. In what follows, I shall assume that 2.a. is the goal.

3. That would make the ideal method of assessment of complexity something like the following:

> Take a representative bunch of taxpayers (or accountants), with skills distributed fairly evenly around the average (but with some highly skilled and some of below average skill), set each of them some hypothetical circumstances, and see how long it takes them to work out the tax due.

> This would not be practical. You would need a big sample, in order to get a decent distribution of levels of skill and to allow each member of a set of hypothetical circumstances to be given to several participants. But I think the idea would still be worth keeping in mind, when developing a practical way to measure complexity.

4. We can start by looking at the list of criteria on page 5 of the document.

4.a. On legislative (and guidance) complexity, do numbers of sections, paragraphs, pages, reliefs and finance acts matter in themselves? A taxpayer will presumably start at the beginning of a chunk of legislation or guidance, under the impression that it might apply to him or her. If the legislation or guidance is well-structured, it will have some navigational aids at the beginning, such as "This is for individuals with rental income, so if you are not like that, stop reading now", and "Only read sections … if you have borrowed money to buy property that you are renting out". A lot of credit ought to be given for that sort of thing. The relevant measure of length can then be the length of the legislation that the taxpayer will actually read, not the total length. Different taxpayers would read different sections, but one could still conclude that the average amount to read was a lot less than the total. (Sometimes the credit obtainable in this way is limited: "This is about companies with loan relationships" leaves you with quite a bit to read, in order to discover what counts as a loan relationship.)

4.b. Some sort of readability measure is needed. Sometimes, a paragraph of pre-re-write legislation can be a great deal harder to understand than a couple of pages of post-re-write legislation. It may be that an established index is as good a tool as any. But it is worth thinking about what counts as having read and understood a piece of legislation. After reading, you should be able to take the right data from the taxpayer's records, and manipulate those data to calculate an amount of tax. You essentially have to extract a mathematical function from the text, and be in a position to use it. This is rather different from reading a few pages of Plato and working out why he wasn't keen on democracy.

4.c. Number of taxpayers is not about complexity, although it can be a valuable guide to where to direct OTS resources. Ability of taxpayers determines the significance of complexity rather than its level, although this does matter, in that legislation X may be simpler than legislation Y, but still more worthy of OTS's attention because X is harder work for X's users than Y is for Y's users. Avoidance risk is not a form of complexity, although it may be a cause of complexity. Cost of compliance and HMRC operating costs may be effects of complexity, and therefore indicators of complexity, but they are not themselves forms of complexity, and they can be effects of factors other than complexity of legislation. If complexity of legislation is what is to be measured, I would suggest stripping out all of these extraneous criteria. If the goal is a measure of the order of priority of tasks for the OTS, by all means keep them in, but then I suggest indicating that this, and not complexity of legislation, is what you seek to measure.

5. The decisions to give scores out of 5, to use a quintiles basis for legislative complexity, and to use a weighted average will obviously have some effect on the final picture presented. My guess – and it is only a guess – is that values will not be of interest, but that orders may be interesting, in the following sense.

5.a. The fact that arbitrary decisions about the mathematical form of the computation need to be made means that we should not attach any significance to the numerical values that are obtained, and we should attach very little significance to the ratios of values (for example, legislation Z scores complexity that is twice as high as the simple legislation W, but legislation V scores only a tiny bit higher than Z). Different, but equally justifiable, arbitrary decisions would have produced different absolute values, and different ratios of values.

5.b. The order of complexity score (W to Z to V in the above example) should be more robust, that is, less likely to change if we were to make different arbitrary decisions about the mathematical form of computation, especially when values are far apart. (This is the minimal sense in which interest might attach to ratios of values.) Thus we could be pretty confident that W was less complex than Z, but much less sure that Z was less complex than V. We can only expect to get a partial ordering out of this exercise, not a total ordering.

5.c. The idea that order should be robust is subject to an important caveat. On page 7, we read "The weightings for each criteria are different and can be adjusted based

on the views or objectives of the user". If different sets of weights are used in relation to different pieces of legislation, nothing can be established about whether one of those pieces is more complex than another. The last sentence on page 10, "Using a single number has the advantage that it is easy to rank across the tax legislation and by adjusting the underlying weights (see above) the final index figure can emphasise a particular factor", can be read in a way that would mean that it missed this point rather badly.

5.d. More generally, I really don't think this permission to adjust weights should be given, unless the weights are established in advance of any work being done, by reference to criteria that can be justified, like the type of taxpayer. Otherwise, there will be great scope to arrange to get to the desired answer. There would be no need to have equal weights (pages 9-10) to eliminate this dangerous discretion, only fixed weights.

5.e. Such concerns do not, of course, matter much if the index only gets used to stimulate debate (page 9). But if it gets used to confirm gut feeling (page 9), one is on the road to thinking of it as proving something. The next thing we know, it will get used in political speeches to claim that governments are achieving great things. It is therefore essential that something on the lines of "This is unscientific and is merely fit to stimulate debate" is said on every occasion of use of the index.

5.f. I am unnerved by the suggestion on page 10 that this index might come to acquire robustness. It is in general true that the more people you ask, the more likely you are to reach the right decision, so long as each person has a better than 0.5 probability of giving the right answer (Condorcet jury theorem), but that assumes you have a well-defined question to ask, and I am not sure we would be able to reduce either the question, "By what method should the complexity of this legislation be assessed?", or the question "What is the complexity of this legislation?", to a set of well-defined questions.

Richard Baron

Institute of Directors

December 2012