ESRC Microsimulation Seminar Series

Seminar 3: 'Moving beyond tax-benefit and demographic modelling'
Leeds, 2nd July 2009

Microsimulation in a cold climate
David Bell

[EDITED TRANSCRIPT]

OK, this is particularly apt for today. OK I'm going to talk to you for a little while about the work I've done on microsimulation modelling over the last year or two. It's sort of a cottage industry. When available I've had a part of a part of a research assistant and it's a model that's kind of grown like topsy, that like Ruth was saying you know, you don't get large chunks of money. So a lot of what I've done is to respond to particular policy issues that have come up, mostly they're Scotland specific, not entirely, and develop my models. So it looks like a lot of bits at the moment and it needs perhaps joining together, I'll explain a little bit about that later. I should say that my, in one of my previous existences I was at the Macro Modelling Bureau at the University of Warwick which some of you may have heard of and we used to look at the Treasury model and the Bank of England model and City University in, CUBS, City University and Business School and so on. And from that experience I've developed a very healthy scepticism about modelling and what, how people could manipulate models to give them what they first thought of. And you know as I go along explain that really some of the things that I'm doing, particularly for policy makers are just illustrations, they're not, they shouldn't be treated in any serious sense as forecast because there are just so many omitted variables kind of analysis that I'm doing that in that sense what you can do is tell a good story and it can be very useful to tell a good story and not necessarily disingenuous.

OK well I've called, we've called it OPERA and in some ways it's very similar to what Ruth has been talking about. Why, what was the genesis? Well really because as Ruth mentioned the situation in Scotland is a bit different from that in the rest of the UK in terms of the care system. And as a result of getting little bits of money the design of the model has been reactive rather than proactive. So I did not have any grand design like MAC 2030 in my head when I started off, I just did things as they came along and it, actually there were some slightly tangential movements along the way, some of which take me right back into times benefit modelling, not beyond times modelling at all and I'll talk a little bit about those.

So what I'm going to do is talk about OPERA itself and how I've put it together, and then talk about the applications and in each case I've actually been dealing with policy makers, policy makers who haven't got a clue about modelling, haven't got the faintest idea, but you have to be able to explain to them how these things work, why microsimulation models provide added value, but why you should be quite careful about the way you interpret the results.

So the first deviation I had was on local income tax, and Sir Peter Bart who ran the Bank of Scotland and one of the most unpleasant people you could ever possibly meet, he was a banker, yes he was a banker! I dealt with him, I dealt with the Finance Committee of the Parliament on the indexation of social security benefits which Ruth has already touched upon, and then the last two things, talking about modelling, home care costs I did for Audit Scotland and the Welsh Assembly and I'm at the moment working on a satellite model on the costs of dementia, the costs of dementia being estimated by Martin Knapp earlier this year at £17 billion per annum, which is about 1% of GDP, so fairly significant and likely to increase quite considerably in the future. So it's a worthwhile exercise. I've done some work with Alzheimers Scotland but also with NHS on that.

OK so structure, well population, the model that I've got doesn't necessarily have to be a Scottish model because its data set is exactly the same, it's set as most other people use.

Actually one thought occurred to me there just at the end when we were discussing, when Ruth was discussing informal caring, one of the things that isn't very good about the Family Resources Survey is that it doesn't do education very well, and if you did view education a little bit better you might get a handle of this issue on whether more educated children are living further away and you could at least use that as a proxy and that we didn't act on your informal care. We use the Survey of Personal Incomes because the Family Resources Survey in certain parts of income distribution isn't fantastic. And I'm always racking my brain about care homes because a large proportion of the costs of care are incurred in care

homes as we'll see from our last section. But all the surveys in the UK including ELSA are only, their sample frame are private house, is private households, not communal dwellings, so they don't go into jails and they don't go into care homes and any other communal dwellings.

OK so structure, I counted, well the same set of accounting relationships, so it works out taxes outside benefits based on people's income and their characteristics, I hope, and it's completely static, it's non behavioural so there's no elasticities of supply of disposable income changes, entirely non behavioural. And the stuff I've done on home care has been parameterised from a data set, an external survey which I did or I did in conjunction with the Welsh Assembly and stuff on care homes is basically calibrated, you can't say that it is foretelling a story, you cannot say that it is other than meeting a few key totals and being consistent I guess probably with FRS data, it doesn't have any particular novelty although I am working with a data set which has got a lot of health information, links the census, the two censuses to hospitalisations, to cancer, to education records, it's a 5% full sample of the Scottish population for the Scottish longitudinal survey but you need to go through myriads of ethical clearances to get access to that and as at with lunchtime I have to have somebody sitting beside me, physically sitting beside me when I used it. However that will actually, we'll be able to tell at least something about the hospitalisations of people who live in care homes which will be an advance that I don't think has been available thus far.

OK, we had, I started off thinking oh I'll write some great software for this and quickly decided no I wouldn't! So I used a combination of stata and mata which is just the matrix programming language that's associated with stata. And of course once you've signed up to a statistical package like that then you can do all kinds of stuff with distributions, with panel data, with survival, and it's possible to write little routines, so if you have regression results they just get dumped straight into the model. So we put all that together without too much pain, it's also very good on graphics, and although I realised I would be in the company of many geographers today, so I thought I would mention the fact that it does do these kinds of maps although that's not in the release. And there is the proof that how much unemployment has increased, there's one between April 2008 and 2009 and the current rates of unemployment by UK travel to work here, that's just straight down and into stata. And that's for an entirely different purpose but I just thought since I knew I would have to pander to the geographers I would put that in.

OK, now let me talk quickly about local income tax, the Scottish National Party got elected to retain(?) the administration not because of anything to do with nationalism, but because they promised to throw out the council tax, and you can quote me on that, I'm quite happy to stand behind that, because there was a lot of complaints about the local income tax. So what the SNP proposed was a uniform of local income tax throughout Scotland, the other proposed by the Liberal Democrats was a different income tax rate in each local authority. And I went through this with the Bert Commission to look at the distributional and spatial consequences. This is just straight tax benefit modelling, the usual stuff looking at the distribution, the distribution of council tax or the burden of council tax in terms of proportion of net household income hits the third decile hardest because the first and seconds get council tax benefit and therefore have to pay very little, the SNP said it would be 3 pence in the £, our calculation was that it would be 4.3 pence in the £, but the DWP said and by the way you won't get the money back for council tax benefit and that increased it to 5.7 pence in the £. So then as you move, as a result of that initially it's the 8th, 9th and 10th deciles that are paying more but then becomes the 7th, 8th, 9th and 10th and then you're just above the middle once you correctly adjust the income tax rate for the removal of council tax.

That kind of graph is quite powerful, you know people can figure that out and understand what it means, they also can understand winners and losers which is quite a simple thing to calculate, how many households win, how many households lose, and politicians, very important, very very important. And here you see if you look at different types of household, the local income tax is the proportion of household income is the yellow bar, the council tax which seems to have disappeared is the blue bar and what ? results were interested in was property tax. Pensioners are the big gainers, huge huge gainers from having a local income tax because the don't pay income tax effectively but they do pay council tax.

So again quite a powerful message, even though you know there's a lot of assumptions in there, there aren't, this is a static model, there are no elasticities of labour supply; for example, fewer people would work when the income tax rate rose and that would further deplete your income tax revenue.

So, now another simple, I was asked about social security, what was very controversial, I don't know particularly why but I guess it was the back end of last year, was the fact that there was an increase in gap between households which were reliant on benefits and households which were reliant on standard incomes, and that gap had arisen because benefits were indexed to prices and not to earnings and that may or may not be remedied. But what we did was actually just a very simple re-weighting. So we looked at the current population structure based on the FRS and then looked forward and thought, went through the third experiment, what would things be like in 2031? Well we know one thing, I know demography is not destiny but it's more predictable than most, so we re-weighted based on the general registrar's projections of the population, both as a whole and geographically to get stuff like this. So you need initially start with the percentage of the households below the poverty line in 2007, 2006 and then the percent who would be below in 2031 if you kept on indexing benefits simply on prices. Now the, so you get something which is really telling you the glaringly obvious, but it is quite interesting that Edinburgh, this is broken down by Scottish local authority, Edinburgh and Glasgow hardly changed, that's because their age, they don't age so much so they are not going to be, because of the pattern of selected migration within Scotland into the cities, they won't have such a relatively large increase in the proportion of families dependent on benefits than the more rural or smaller city parts of Scotland.
So glaringly obvious but a useful way of describing that to policy makes.

OK now this is home care costs, and this is tricky because I'm touching on the same issue that Ruth has mentioned, so it's people who are receiving care in their own homes, so social worker care assistant is coming along and doing something for them, cooking a meal or whatever they would do. I did this survey in Wales and what became abundantly clear is how skewed the distribution of costs is. So actually it's the case that 10% of the people account for 40% of the costs and 40% of people account for 10% of the costs, and that actually makes microsimulation quite difficult, it's not a neat little normal distribution that you're dealing with, you're dealing with a very skew distribution and OK, yeah … (clicking)

Female question 1 – Can I ask for clarification on something? When you say cost what do you mean? Do you mean?

David Bell – Local authority reported costs.

Female question 1 – So some of that skew is because lots of people pay enough in to use it.

David Bell – No this is not charges, this is not charges, this is actual costs of the charges I have set.

So it means that some people are getting very intensive packages, whereas the average is something like 6 hours a week. And mostly with cost functions economists approach this by modelling the log and given the distribution you can see why that would be an attractive thing to do because certainly it didn't look normal, and if you take the log, if you model the log of the cost you get into issues with simulating that which arise from the non linearity of the logarithmic function, and we played around a lot with this before we got, well I wouldn't say we got a completely satisfactory solution because you could alter one of the parameters and get quite wide variations. Effectively the expected value, if the distribution is at all heterogeneous capacity which it is then you have to take into account that heterogeneous elasticity in your conditional means.

So, given that nevertheless what we were able to show was that those overall costs increase with disability, decrease with age, decrease with presence of informal carer, are unaffected by gender and ethnicity, and do as we were just talking about, vary by local authority, so some local authorities are more generous than others.

So for example here are the average weekly costs of personal care, that's intimate care provided by Welsh local authorities, a selection which covers more than half of the Welsh population, by age group, male and female, more for younger people because younger people tend to have more complex disabilities, older people it's more frailty. Now this is an example of why having slightly different policies in different parts of the UK is quite interesting because the Scottish policy only applies to 65 and over, so there has been a question well should we extend it to those under 65? And that's been resisted partly as a result of information like this because the costs would be quite horrendous, whatever you might argue in terms of social justice, the costs would be quite dramatic.

OK we have our what's called in index of resource need that's used in Scotland to classify disability and each ? are people with, probably with dementia and all kinds of care needs. And so mild disability for males and females is attracting an

average like say £50 a week whereas we're four times that, and that's on average for people with complex needs, again highly skewed distribution.

And then this is your charges, so this is the charges by disability level, it's going up to £60 a week on average. So on average the Welsh population were being charged a third of the costs of more intensive forms of care, I think there should be a tenth one here unless I've numbered them wrongly. But that gives you an illustration. And of course this is quite a good way to take a lesson from one part of the UK and apply it to a different part of the UK because the Scottish legislation is about 2 lines and it says local authorities are prevented from charging for personal care, that's it. Well this must be what they're giving up. So that's one way of looking at the costings or applying them across to Scotland.

OK, so this is what we did, we estimated the determinants of the costs of care using the Welsh data set and then did some regressions with the Family Resources Survey and matched really on disability rather than on a more complex matching system and you know gave those with the most complex disabilities as far as we could assess from their ? (25.00). We would bring then into the equivalent of the Scottish policy and to get any stability we're effectively doing Monte Carlos(?) to get, because of that skewness problem, and the re-weight of the results as well.
Right I'll come back, if anybody wants to ask me more questions about that they're welcome to.

Right where we are now is actually on the model of dementia costs. At the moment it's an entirely, it's just a satellite, so we're not thinking of just private households, and what we're using is simply information on life expectancy, dementia prevalence, duration and costs. And we have a competing risks model, so the risk of getting dementia and the risk of dying from other causes is there, it's fairly simple at the moment. What you can do with Stata and it's quite easy to do little time aggregation exercises. So what we're doing is projecting forward the population of older people in months, because a lot of these costs is easier to think of them in months and spells don't follow years.

So we use hazard to model the months of life expectancy after age 65, we approximate as well as we can the characteristics of the current population. So here's, of course what you can do, because we have data on a life tables by deprivation indices by deciles of deprivation, monthly ten(?) and there's the famous story of you take the train from Bears(?) down to, I always forget where it is but it's somewhere in the East End and life expectancy changes by more than 12 years over that train journey. So actually the life expectancy for men for the first decile I think is 67 and in the top is 79. And the different, and it's particularly men that are affected in the bottom decile, the gap between men and women is much narrower at the top. So we're able to get the life expectancy on the one hand and then what we've got to think about is prevalence of dementia and it seems, I think I took this from the LSC study of, there's not, I have wandered my way through quite a lot of the medical literature looking, well there's some consistency maybe about prevalence but then length of the disease. So you've got to, so you have someone who is, who if they're not struck by dementia would last for X months, you then have a risk all the time they might catch dementia, how long do they last thereafter? And the, I'm happy to be corrected but the mean that I have taken from the literature is around 6 years and you know that's, I may not have got the right bits and pieces out of the literature to do that.

So now what, why, well it might seen obvious why I'm doing this but I'm involved with the NHS who want to shift the balance of care and by the balance of care what I mean is they want to shift the balance from the care home sector and into home care. Scotland's done that in quite a big way already and I'm going, I'm being asked to help managers understand this process and to use something like a model like this to help them understand what, what are the processes that are going on. So it's not, I'm not really projecting I'm just trying to help people think through what are the issues involved. So you know what if the prevalence changes by age group because some new drug has been found, you know, the company that finds an effective drug for dementia will become very very rich. So if you could delay the onset, what if the cost structure changes, so the balance of costs between care homes and home care changes, but also with part of labour cost or capital cost change.

So going back to my 1st and 10th deciles, here's the 1st decile and here's the 10th decile, because men don't live very long at all in the first 1st decile their incidence of dementia is really quite low. Now you might say, well it would be true to say I don't have separate evidence on dementia incidence for people in the first decile, you could say well a lot of this is alcoholism, so there might be an issue about prevalence there. For men and women in the 10th decile is broadly in terms of, broadly the same but because more women last, live longer, overall more women are affected by dementia.

And so what we do is to simulate a population going forward, hazards for a set of people starting back in 1981 and then taking them forward in a way that's consistent with life expectancies, each year going forward to 2030, well you've got to start further back because some of the people in 1980, who were 65 in 1981 will only be developing dementia now. And split the periods into severe, moderate and mild with the highest costs associated with severe, intermediate costs of moderate and so on, and just look at changes in prevalence. What I've actually done is to change the age, drop the age specific prevalences by 5% and that gives you lower, that comes in both cases. But I haven't really calibrated all of this yet, so I do have, taken from the study of Joseph Rowntree or Lennon Fuisson (?) did for Joseph Rowntree in terms of capital costs, the staff costs, the other non staff costs associated with care homes and then the care at home costs which are drawn from the Welsh stuff, care assistants, nurses and management costs for home care, HCS(?) home care built up possible changes in costs associated with again decreased prevalence. And I have just assumed an allocation between home care and the care homes.

So that's, I guess it's true to say that the model is not specific to dementia and I could run, could have competing risks for all kinds of stuff. But what I would welcome any help on is to integrate effectively what is a shell of a model which meets some very broad population parameters into the main FRS data sets. And also this issue of, I keep going on about this, but how to deal with those people who are care home residents who will just always have to stay in the satellite because we don't have the socioeconomic information on them.

So that's the dementia model. And my overall conclusion, calibration like this, when you're talking to policy makers does help explain the obvious, and it's no more powerful than the data on which it's built and therefore, because politicians get terribly impressed when equations are involved, you should not overplay what you've got and work with them very carefully. OK. Thank you.

QUESTIONS

Male question 1 –On the costed chart for Wales, the one that you said reinforced the policy decisions of Scotland, the costs for ??, well the chart was showing costs per person receiving care, presumably that ???

David Bell – Yeah I know what you're going to ask.

Male question 1 – There's far fewer younger people receiving care …

David Bell – There is about a sixth, OK there's a sixth and their costs were twice as high on average, so you were going to add a third to the total costs, and that's why basically its unacceptable and indeed the Welsh saw this and decided not to go ahead with it.

Male question 1 – But it was the average across those receiving care, not the average across the whole population for that age group. Because the older groups have got far more people receiving care, so the total sum costs would be greater than the ???

David Bell – Well all that I had were the clients of 4 local authorities and those were selected by the local authorities to receive care. There were a sixth as many young people, people under 65 as there were over 65, they cost twice as much as the over 65s. So from the local authority perspective this group constituted a third of the costs. Now clearly going forward with smaller cohorts of younger people, that might diminish relative to the older group because, it would be an intermediate issue to think through, but through time I guess relatively the younger ones are going to cost less, if the same cost ratios continue to apply.

Male question 2 – (38.04) I was just thinking about Downs Syndrome really

David Bell – Yeah absolutely, very important.

Male question 3 –The council tax versus local tax? Who were the losers in the poor population under the local tax scenario, what kind of characteristics …

David Bell – Well multi owner households with higher incomes because they'd both be paying this tax. So you know it's Mr and Mrs Middle Scotland who reacted badly. So those groups would lose, the dual earner households, the ones who were on state pensions were most clearly …

Male question 3 – Because I saw there was a very small proportion in the pool lost under the …

David Bell – Oh well yes I see it all depended on the rate, if the rate was only 3 pence then you know the dual earner households didn't lose to the same extent. But once you put up, the rate up to 5.7 which was the correct replacement rate, then these households would face considerably higher income tax bills.

Male question 3 – And on your penultimate slide you said you need information on socioeconomic demand for home care? Do you have any information on them other than age or …?

David Bell – Not that much.

Female question 1 – I mean that we deal with that is, because we're just using the FRS information, the re-weighting is the way we deal with it and that, the weights come from the PSSI new model and they have limited information from one off surveys which tell you something about, mainly about whether people were previously owner/occupiers before they were in …

David Bell – Yeah, I mean in a sense they should represent partly the population.

Female question 1 – But it's kind of, it's patchy information which ???

Male question 3 – So you've got nothing on them other than how old they are?

David Bell – Well I've got these data that I'm just starting to investigate on the actual health costs ?? and we may need spatial information which might, you might take as a broad indicator, so they come from one of these multiple agent(?) deprived areas.

Male question 3 - more geographic ?

David Bell – Yeah it would be approximate.

Male question 4 – You mentioned about integrating the dementia model with the ?, I mean in one sense if they don't actually fit together then you may be better off you know just saying that and like you attach some static model to a dynamic model and I mean people don't, I mean is it for intellectual reasons you want to integrate or is it for some kind of…

David Bell – Yeah, I mean the most obvious thing is to be able to say something distributional, so you then need a characteristic or several characteristics that would allow you to hook into, approximately hook into something that ? and whether that is just a step too far anyway, you know it could, if that were feasible and believable, you might get some nice stories about distributions thereafter. But I'm not sure these, that condition holds.

Male question 4 –Most people say that dementia is a distribution issue in a way, it's one of these issues the criteria of treating dementia would not necessarily, distribution effects would not be a major part of that discussion.

David Bell – Only for, it depends to what extent charges might apply. In the moderate phase it might be the case that the local authority would expect some means tested, means tested in respect of …

Female question 1 –It's the classic result of what's given of diagnostic inequity in that because there are no health, available health care treatments that the NHS will offer limited what care they do get is open to means test. So yes, but then you need to know some things about ?? incidence dementia which we you know …

David Bell – Yeah, although people who work longer are less ?? dementia.

Female question 2 – Could I ask just for clarification, I think it's your chart of the indexation ? effects which you have classified by Scottish level of ? is that based on FRS data or are you, well my question is are you able to analyse the economic ? Scottish ??

David Bell – No there were time, not the last, well sorry one thing I should have said is the Scottish sample is boosted. You can, I didn't get for last year but for the previous year when I did this you could get it broken down by 6 areas within Scotland and then it's based on imputation within those 6 areas. So it's not …

Female question 2 – I mean I think my question is to what extent that variation is capturing variations in receipt of benefits within the ?? areas.

David Bell – Well we do know the, there's the thing called the Scottish Council Survey which is representative of the local authority ? so we can get quite good data on incomes and what people, which we've used. We also get quite good data on what people say they get in benefits, I must admit I haven't, I didn't use that for ? but they do, that's a 50,000 sample within Scotland and they're asked what benefits they receive.

[END OF RECORDING]