ESRC Microsimulation Seminar Series

Seminar 4: 'Bridging the Gaps; Setting the Agenda'
University of Sussex, 11th September 2009

Microsimulation Modelling in the DWP – Lessons Learnt and Future Challenges
Chris Drane

[EDITED TRANSCRIPT]

OK, I'll start by introducing myself, my name's Chris Drane and I head up the Model Development Unit at the DWP. What I'm going to talk to you about this morning is actually not the detail of our tax and benefit simulation models or other microsimulation models, but rather how we go about doing modelling within DWP. I'm going to talk a little bit about some of the things that work well and the way that we do it, and some of the things that really don't work quite as well. I'll also touch briefly on what are some of the main modelling challenges going forwards.

I mean the main thing to learn about how we do microsimulation modelling at DWP which David touched on earlier is that we have a team that is dedicated to model development - so that's a team of around 20 analysts of different professions, we all sit together in the same office typically and what we do is we are a ring fenced bit of resource for the department if you like that focuses on development, maintenance, and support primarily. We then provide the models on the desk tops of professional analysts across DWP. Those analysts then typically go on and produce the analysis that ministers want and need.

Most of the work we do is in-house with this team, sometimes we contract in the expertise, particularly where there are particularly knotty econometric issues and we need some real econometric expertise to deal with it, labour supply being a prime example of that where there's external expertise that can really help us along.

We do do a little bit of analysis using the models ourselves, certainly in the last couple of years, and that's primarily in support of the department's Strategy Unit, who are a team that organisationally are pretty close to us and who don't have much analytical resource themselves, so they tend to call on us directly to do bits of modelling.

I'll just run you through some of the history of model development in DWP and some of the key points in time. Pre 1999 the main modelling tools that we had within the department, what was then the Department of Social Security, were a static tax benefit microsimulation model, just as Mike, Dave and Doug have described to you, known as the Policy Simulation Model or the PSM for short. We had a model called PENSIM which was a long run dynamic microsimulation model of pensioner incomes that had been developed externally and was used within the department. Those models were maintained by policy analysts and when I say policy analysts what I really mean are people who have lots of other stuff to do other than model development, so people who are answering parliamentary questions, people who are doing briefing for ministers and doing a whole range of other analytical work and policy related work, who are trying to develop and maintain the models and provide support for them, as well as doing all the other stuff.

What we typically found in that situation was that the PSM and PENSIM were not receiving sufficient attention in terms of model development and they were falling into disrepair. The rationale for that is pretty straightforward - if you've got a parliamentary question to answer or an urgent bit of briefing or analysis for ministers, that is always going to take priority over something like going and updating your base dataset for PENSIM or spending time updating your econometric regressions or similar work.

So the decision was made in 1999 to ring fence the resource and to create the Model Development Unit. And the first jobs of that unit which then was about 10 to 12 people was to convert the Policy Simulation Model from an in-house language into SAS so it was a bit more accessible to people around the department, and to begin the development of PENSIM2. The first PENSIM had some limitations in terms of the data on which it was based and just how maintainable it was, and so a large project was started to bring into place a new long run dynamic microsimulation model of pension incomes which was a bit more future proofed.

The next key milestone you'll notice is a good 4 years later and that was the delivery of a first prototype of PENSIM2. That was a really big project within the unit and it took up a lot of time for a lot of people, getting that model developed and we've certainly learnt a lot of lessons from it which I shall run through later, but it really was a tough job delivering the first version of PENSIM2.

A year later another really key milestone for us. Through the course of developing PENSIM2 we developed our own in-house software for running dynamic simulation models called GENESIS. GENESIS was very much the brainchild of Sally Edwards my colleague who is sitting in the front there who will be able to tell you lots more about it later on, but it was a very flexible and quite a user friendly programming language in which to develop dynamic microsimulation. And what we were able to do was de-couple it from PENSIM2 and start developing new dynamic microsimulation models, using that sort of software, that toolkit.

2005, the first real use of PENSIM2 by the Turner Commission, so a big moment, actually the first use of a dynamic simulation model to influence policy development, so quite a big moment for the department, but also quite a big moment for dynamic microsimulation within the UK as well. And also organisationally important for us - Strategy Directorate was formed which was basically the grouping of us alongside colleagues from the Strategy Unit which has slightly changed the way in which we work.

Since 2005 we've greatly expanded our modelling capacity within the department, so several more dynamic microsimulation models, particularly on the benefit forecasting side, for medium term forecasting over a 5 to 10 year time horizon, so forecasting retirement pension, incapacity benefit, disability living allowance; more recently developing an integrated benefit forecasting model for the working age benefits that we administer; and also recently developing a dynamic microsimulation model based on the Labour Force Survey to help project employment trends - that's a model which we're currently in the process of validating and testing.

We've made enhancements to our static model as well, a 2020 version of the policy simulation model is not dissimilar to the work that's gone on with the IFS and at Essex which allows us to look at the income distribution in 2020 and see how poverty trends might evolve.

A labour supply model, the Employment and Hours Model that sits alongside the PSM as well, and that broader role of doing a little bit of running the models to help support our Strategy Unit.

In terms of the sort of successes and real achievements that we've had, I think it's important you know from our point of view, success for us is typically judged not only in having a great model that's technically correct and does really good things, but also in actually seeing those models being used to influence policy decisions. Really what we're about is trying to help improve the quality of the key decisions that are made in the department or across Govt more widely. So some really big successes there. As I mentioned PENSIM2 was used to underpin the work of the Turner Commission and a lot of the subsequent Govt work on pensions reform. It's now established as the key tool for long run pensions analysis in Govt and enabled analysis of the decisions that went around Turner and the Govt's reform programme, allowed some analysis of those decisions that just otherwise wouldn't have been possible. So PENSIM2 really supported the thinking around pensions reform and improved the analytical capacity for pensions modelling in the UK.

The Policy Simulation Model continues to be used very widely across the department, that's a model that's like IGOTM in Treasury and like TAXBEN at the IFS that has been passed down for many years in various forms, we currently have I think probably just over 40 active users within DWP who have got the model on their desktops and are actively using it on a weekly basis to analyse different policy issues. PENSIM2 the user base I think is probably around 15 to 20 people actively using it. So both the models are really actively contributing to decisions and you know it's quite rewarding to see that happen on a regular basis, decisions influenced by your modelling.

I mentioned some of the enhancements to the PSM. The static ageing of the PSM to look at the world in 2020 has had a really big influence on thinking around strategies to alleviate pensioner poverty and to alleviate child poverty and is continuing to do so, as is the work on labour supply effects, when considering the work incentive effects of policies. And in terms of pure modelling rather than influence on decisions, the building of INFORM, our integrated benefit forecasting model and the dynamic LFS model has been particularly pleasing because we've done them to quite demanding timescales. Whereas PENSIM2 took us 4 years to get a prototype up and running, and another couple of years before it

was actually used in policy analysis, INFORM because we've got the GENESIS software took us about a year, the LFS model is still being tested but we got a prototype up and running in just under a year - both with much smaller teams - so we're becoming quicker at doing this sort of modelling and that's been a real achievement.

Over the course of a decade of having a Model Development Unit we've learnt quite a lot about things that work well within our organisation and things that work less well. For us the ring fencing of resource to do model development has worked really well and it's been really important in allowing us to build up a really powerful toolkit of models and really quite a substantial modelling capacity. Doing that has required significant investment and significant support and a willingness to take risks from senior staff in the department which has really been to their credit to say 'OK we will fund modelling projects, we know that they might not come off but we'll give you the time and the space to see if you can improve things'. So a real commitment to investing in analytical capacity from senior staff in DWP.

I think we've got quite a structured approach to model development which I think has served us well and that's both in terms of making sure that we set up quite formal governance structures for modelling projects e.g.steering groups, groups for customers to give feedback within the department. We also adopt quite formal project management procedures in running our model development and we have quite an organised way of making changes to the models as well in terms of when users spot things they want to see improved or bugs that they spot in the modelling code we have quite an organised way of logging that and prioritising making the changes, which is important when you've got a large user base. That really helps us to manage priorities and manage the development of the models.

The way in which we staff the team I think has provided us with some good disciplining factors. Typically the way in which the junior staff in particular within the Model Development Unit come in is, quite often they're fairly recent graduates but they typically work within the unit for one, maybe two years and then move on to other posts in DWP, so there's quite a rapid turnover of staff. Now that has its pitfalls as I'll mention in a minute, but the upside of it is that it does mean that we need to make our models relatively easy to maintain, if you don't people just can't pick it up in time. So it was one of the key drivers behind building the GENESIS software in the first place and it continues to be a disciplining factor in the way in which we code and the way in which we try to document.

The desktop user model is also a good disciplining factor because it means that models need to be relatively easy to use, they need to have if not front ends, quite transparent coding, there needs to be a good user support structure in place, you know it just helps us think about how we're developing the models which is good for the long term maintainability and the likelihood of us keeping them going in the long term.

A really big thing for us and a thing that we've done really well is building a large customer base, and that's important on a couple of counts. Firstly getting lots of people using our models helps make sure that they are influencing more decisions and they're feeding into the policy process as much as possible, and it's also sort of a bit of a virtuous cycle, the more people you get using the models, the more demand there is to actually make the models better and improve them. And actually going out and marketing the models and the capacity of models therefore becomes a key part of what we do to make sure that people are aware of them and to make sure that people are encouraged to use them, and certainly what takes up quite a significant proportion of my time is actually going out to people around the department and making sure they know what's available and making sure that they're using it.

A couple of points, important things that I think are key in having a modelling function and developing models that work are related to what is going on outside the department. I know Mike touched on this and I absolutely agree with him. We've certainly benefited from an enormous amount of support from the academic community and the international community in developing the models, and without support from the SAGE team at the LSE and Southampton, from the IFS, from people working at DYNACAN in Canada, NATSEM, people working in the US, we wouldn't have got nearly as far. We really drew on a lot of expertise that was out there, particularly in the development of PENSIM2. And having challenge from outside the department is important. As Mike mentioned, an important factor in driving up the demand for high quality internal analysis is high quality external analysis. So one of the reasons and one of the key factors in allowing us to get investment in modelling capacity within DWP is knowing that we have the likes of the IFS and the Pensions Policy Institute and others doing high quality work outside. Also what we've really found is that getting external people in to actually cast an eye over our models really helps the credibility of them internally. Some classic examples are on the PENSIM2 side which was a model that we really needed to work hard to establish internal credibility for. Some of Mike's colleagues at the IFS came in and audited the methodology and a couple of years further down the track we got a senior

member of the team who worked at the US Congressional Budget Office to come over and cast an eye over the model and write a report on it to senior people in the department. Really, really important in giving decision makers the confidence that the modelling can be trusted to say sensible things.

Now of course over a ten year period there are plenty of things that we learnt didn't work so well as well. The first lesson we learnt I think is that the demand for analysis is such that developing, maintaining and supporting models for use in a department like DWP is a full time job. It was one of the reasons we needed to put in place a Model Development Unit in the first place was that there just wasn't the resource available to develop the models any other way. And it's something we found as well when we've had experience of contracting out the development of models which has sometimes been a challenge. What we've tended to find out is when we've contracted out the development of models, we've often got back excellent bits of modelling but without having that expertise in house on a day to day basis to support people who are using the models, it's very difficult to provide adequate support. So we've tended to move more towards in house development where we can, simply to ensure we have sufficient expertise on hand to provide that support. It's also quite difficult to find high quality external researchers who are willing to provide the sort of permanent, daily support services that we need..

I've spelled out some of the benefits of our way of working with the ring fenced bit of resource and with the way in which we have our desk top user model and the turnover of staff, but there are downsides to that way of working as well. You know an obvious one is that with a high turnover of staff you lose a lot of expertise and as the 3 members of my team here will testify, it's quite soul destroying to spend 18 months or a year getting a young analyst up to speed on the model and just at the moment where they're really doing excellent things with it, they're whisked off to another post and you've got to train up somebody who is brand new and new to the model as well. And even with models that are as easy to use and easy to maintain as we can make them, there's no getting away from the fact that some of these models are complex and it is difficult to get up to speed on them quickly.

Building up the broad base of expert users is a hard slog. It's difficult, particularly with new models, to get people to engage with them when they're complex and they're hard to get your head around, and it's a constant battle to keep making sure we have a good pool of expert users, when they're turning over quite frequently as well.

You also find when you expose microsimulation models to a wide range of people, both users and developers, that they're not for everybody. For many of the analysts who have had to work on the models, either maintaining them or using them, I think they've found that it's not what they want from a professional career,, it can feel like quite a back room existence (?20.06) relative to other Govt analyst jobs. you know sitting there in front of a model all day, every day, some people prefer to be getting out and interacting with ministers more for example. So we found that out, that model development is not for everybody.

In terms of the modelling itself, key things that I think we picked up was that IT can be a real barrier to successful development of models. Slow run times and the problems with the software which certainly beset a lot of the early development of PENSIM2, - we were still sort of developing the software as we were developing the model - can really kill the pace of development.

Another challenge in terms of the model development is dealing with uncertainty rather than risk. This is really an issue of how you explain to senior people who have funded you that you're quite unsure when you're going to finish developing a model because sometimes you really don't know, and it's a real problem particularly we've found with the dynamic microsimulation models where basically everything depends on everything else, so when you get into validating them you start fixing things and all you do is just uncover other bugs elsewhere, and until you've got everything working in a reasonably sensible way, you haven't got a sensible overall picture. And trying to explain that to senior officials in the department and explain that you think you'll have a model ready by the end of September but it could be October, it could be December, it depends how it goes, is a hard sell.

And a really key lesson that we learnt is, and one of the reasons that we've got better at developing models I think is to start simple. When we started building PENSIM2 we started out really going for a complex design straight awayl, and we have really found that it is better to start with the simplest thing that you can get away with and build it up over time, it helps develop customer confidence and it helps keep up the morale of the team, and it just helps you in terms of making

4

sure the models are transparent and sensible as you go along. It helps with the validation. So start simple is absolutely the biggest lesson.

In terms of the future directions and challenges, we've got a pretty powerful toolkit now for analysing policy reforms and for forecasting caseloads and expenditure on our benefits. A key thing going forward is developing some of that modelling - and we're getting there with our INFORM model - to understand more about the delivery of services as well as policy. And I think that opens up in particular the field of spatial microsimulation more, it's not something we're actively developing at the moment but I think that's a potential next step in terms of where one looks for the models to go in trying to help apply some of these techniques to making delivery better.

I think looking hard at the extent to which macro economic and micro economic models link up is also important. One of the things that using these models through a downturn shows is that there is a risk that they are developed and used without full consideration of the macroeconomy, particularly when there are, with some of the dynamic models especially, a whole range of assumptions that go into them. It is quite difficult to think through in your head how all those assumptions should change in different macro economic scenarios. So one question and one thing that we need to consider is to what extent you need some automated macro model that constrains the microsimulation models. We could potentially use models elsewhere in Govt or external to do it, or develop something simple internally, but I think that's something that microsimulation modellers more generally need to be thinking about.

And keeping pushing on the behavioural modelling side of things, as computing power increases there are more opportunities to do structural rather than reduced form (?24.30) modelling and you're seeing it increasingly with places like NIESR and with the labour supply modelling that we're doing and Treasury are doing as well.

I think there are always challenges that we face in terms of cross Govt modelling, making sure that the likes of myself and David and Doug are talking to each other. We do, which is good! But also making sure that departments across Govt make use of the modelling capacity we've already got, and there are some real opportunities there for augmenting things like PENSIM2 to look at issues like generational accounting or to look at issues of long term care. And so one of the things that we're keen to do is to be encouraging other Govt departments to put in a bit of their time, not to reinvent models but to work with what we've already got.

There'll almost certainly be increased pressure on our resources over the next few years and I think that probably goes for anybody working in Govt, and that's something that we're going to have to manage carefully.

And increased scrutiny of results as well, not only to justify the resources that have been put in, but a model like PENSIM2 which is a model which runs out to 2100, one can scrutinise what comes out of the model in terms of the early years of the modelling, but the only thing you can compare the results to 2100 against are other people's models, and there are reasons why we don't always trust those models as well, that's why we built PENSIM2! Actually we've moved more in the last few years into some modelling that will have results that can be more immediately scrutinised. So our forecasting models, you know the results are very carefully scrutinised by colleagues at the Treasury who want to check that we're getting our forecasts of benefit expenditure right. Our labour force modelling as well which runs out to 2020 is going to be much more carefully looked at, as is some of the work around poverty projections that ourselves and the IFS have done. So there is rightly increased scrutiny of results.

I will finish there. Any questions?

QUESTIONS

Questioner 1: You mentioned validation, what's your general strategy for validating dynamic models?

Chris Drane: I mean it has to be a fairly multi faceted approach because there are lots of things to validate with dynamic models. I mean we try to spend quite a lot of time walking through the code, it's very easy when coding up dynamic models just to make simple coding errors that are difficult to spot. So we spend a lot of time doing that. Then a lot of it is once you're convinced that mechanically it's turning right and you're processing things in the right way, you know we tend to compare the models against whatever else is out there in terms of forecasting or projection models. So you know the classic example for, a model like PENSIM2 was comparing against the long run State Pension Forecasting Model that was

produced by the Govt Actuaries Department. So we did a lot of comparison with much simpler models to check that we were in the right ball park. On some of the models we have a base data that is not the most recent year, so you can actually compare the early years of simulation against actual data as well. So we have a range of things, so you know comparing it to other models and other projections where we can, some general sense checking as well, we try and get as many people within the department to have a look at the results of models as we can.

Questioner 1: Do you run them all many times? How many times do you run the model?

Chris Drane: We probably don't do as much analysis of the stochastic variation of the models as I think we would like to, and a lot of that is down simply to the speed at which some of them run. Perhaps Howard and Sally as owners of PENSIM2 might say a bit more about it but we run it several times to make sure that the random numbers between model runs don't lead you to vastly different results, and typically we found that they tend not to, so the random variation of the models doesn't have a huge effect particularly on very high level outputs. We have the capacity in all our dynamic models to fix the random number as well which allows us to do a bit more careful diagnostic work on why particular results arise.

Paul Williamson: – Could you tell me a bit more about the INFORM?(

Chris Drane: OK, I mean INFORM is, typically how we worked in department for forecasting benefit caseloads and expenditure has been a lot of separate Excel based or indeed microsimulation models looking at individual benefits. INFORM makes use of something called the Work and Pensions Longitudinal Study which is basically 100% benefit data, and it's a database that basically links up all the individual benefit administrative datasets. INFORM at the moment forecasts caseloads for working age benefits, but it's an integrated model so it's trying to capture all the benefits in one model. A couple of advantages of that: one is that it allows you to capture the interactions between different benefits in a more consistent way and hopefully fine tune the benefit forecast. The benefit forecasts were pretty good anyway, so the main gain that one gets from having an integrated model like that is that it will give us a better picture of how the department's customer base is likely to evolve. So for example how the number of people claiming different combinations of benefits is likely to change, are we likely to see an expansion of people claiming multiple benefits? Is it going to be more people who are having less frequent interactions with the department? How they're spread across our different agencies and the like. So it's going to be a really key tool for thinking about our business strategy and how we actually plan our service delivery as well as a forecasting model. And it's a dynamic microsimulation model.

END OF RECORDING