ESRC Microsimulation Seminar Series

Seminar 2: 'Adding behaviour'
London, 6th May 2009

AGENT-BASED MODELLING AND MICROSIMULATION: NE'ER THE TWAIN SHALL MEET?

Edmund Chattoe-Brown

[EDITED TRANSCRIPT]

Right, thank you all very much for inviting me. What I'm going to do, first I'm going to do some apologising in advance. It's always a tricky business trying to compare approaches in general, there's a requirement that you generalise, on the other hand you've got to avoid the 'my Uncle Charlie smoked 80 a day and lived to be 100 therefore smoking is safe' logic. You can probably come up with one paper that goes against some of the things I've said but that doesn't mean the general tendencies that I'm describing are not general tendencies. Nonetheless your mileage may vary as the Americans put it and you may know that your model is proof against some of the concerns that I'm going to raise. What I'm going to do is I'm going to try and raise a number of concerns or questions for discussion and I'm relieved to see that these are live issues in the paper that was just presented, around a simple example of agent based modelling, because I think it's reasonable that at least some of you won't have a clear idea of what is involved in agent based modelling. So I'm going to present a very simple example and I know it isn't in the least bit realistic but it does nonetheless demonstrate the principles and get us thinking about some of the things that we need to worry about. And the question is, and some of you at least will know this model already, I'll do a quick head count because then I can speed up the explanation. How many people know the Thomas Schelling model of spatial segregation? OK that's about half, so that's alright, I'm not going to be boring too many of you, and maybe this will be a different take.

The research question is how do we explain the urban residential segregation between ethnic groups? It's an incredibly simple model, it was virtually done before computers. A picture is worth 1,000 words, here's a simulation written in netlogo, you have 2 types of agents called red and green to avoid racist overtones located on the grid, so everybody has 8 neighbours, however some of those neighbouring sites can be vacant, empty houses if you like. So each person's neighbourhood is defined by the 8 agents around them, or the 8 sites around them, some of which may be empty.

Initially you distribute the agents in the empty sites randomly, every agent decides what to do in the same very simple way. They all have what is called a preferred proportion of neighbours, so a PP of 0.5 means that you want at least half your neighbours to be your own kind, the reason we use fractions is because empty sites are considered not to count, they don't upset you or encourage you, they're just empty sites. If an agent is in a position that satisfies its PP then it doesn't do anything, it stays where it is, it's quite happy. If it's in a position that does not satisfy its PP then it moves to an unoccupied position chosen at random. Clearly this is a ridiculous model of ethnic segregation, there is no housing market and so on and so on and so on, we know all this stuff.

Even with this silly little model however interesting phenomena start to arise. Two questions for you, I won't do this as question and answer, what's the smallest PP, a number between 0 and 1 that will produce clusters? Second question, slightly harder, what happens when the PP is 1? Now depending on who you ask, you may or may not get these two surprising answers. A PP of about 0.3 will produce decent clusters. You might assume that people have to be xenophobic in order to want to live together. In fact that's not the case, people are quite happy to be surrounded by a majority of the other type and clusters would still form. On the other hand if you'd seen this data in the real world and you were a sociologist of a particular flavour or a politician of a particular flavour, might you assume that people were somewhat xenophobic? To what extent do we attribute to the patterns we observe certain kinds of behaviour based on our theoretical preconceptions?

More worrying and more interesting for my discussion, as people get more xenophobic the clustering gets stronger, clusters get more separate, they get tighter, they have less direct contact, they tend to be buffered at the margins by empty sites. But at some point these clusters start to break down and with PP = 1, total xenophobia, the system looks no different from the random starting position. I won't bother to explain that, although if anybody's interested or can't work it out, they can ask me.

But what we have here is a fairly sharply non linear system. If we look at the PP along the bottom between 0 and 1, sorry the scaling's a bit rubbish, this is an Excel spreadsheet, we can see that up until about 75% nobody's unhappy, they can all get where they want to go. Similarly the level of clustering goes up to that point, people are in similar clusters with people like them, until suddenly at about 75% there's a regime shift. The percentage of similarities starts to drop sharply down again, the percentage of people who can't satisfy themselves rockets up to 100%. What's happened here is that people's individual preferences have suddenly become incompatible at the global level, it is no longer possible for them to all want what they want and achieve it. It's a system level shift.

So even for this very simple system where we're just looking at the same behaviour, the same PP for everyone, the same very stripped down domain, we've got non linearity emerging and that will turn out to be important for what I'm going to talk about now.

Quick aside about this. It's a silly model, I know it's a silly model but it can still tell us something about data. If we were going to stop this being a silly model, what would we have to do? Well, the individual level model would probably collect data by what we recognise as qualitative methods, you might actually go and talk to people in their houses, you might mooch around their neighbourhoods and see what they look like and what they felt like, you might conduct experiments where you put people in particular settings to see how they behaved. You'd be creating in this way a testable set of hypotheses, we made hypotheses here about how people made decisions, we assumed they all made decisions in the same way which is pretty strong, but that's nonetheless a set of hypotheses up for test, hopefully based on data, calibrated on data.

On the other hand in aggregate data the cluster patterns for example are likely to be collected quantitatively using household surveys and censuses, using GIS data if that was relevant and available. And the simulated outcome of the individual actions is falsified against similarity between simulated and real data. Does this look anything like the city of New York in residential terms? And then there's the important question about what we mean by like, well plainly this doesn't because it has no streets, it has no city blocks, it has no parks and so on and so forth. But it might if we started off with reasonable micro assumptions, we could then generate interesting patterns of clustering and see whether they looked anything like the real thing.

Now these are things I shall claim, I shall try and justify them as we go on. Interesting phenomena, interesting properties of this model, it has no fiddle factors, it's completely defined, there is no noise, no uncertainty, there is no curve fitting, there is no use of correlations, it is all completely causal and deterministic. Now that may turn out to be a drawback, we'll get to that, but it is the case this model is completely specified and causal. It contains no theory constructs, the nearest you get to a theory construct is PP and even that's describable in everyday terms. To what extent does somebody feel unhappy depending on the characters of their neighbours? So we've got no discount rates, we've got no cognitive dissonance, we've got none of these theoretical constructs. There's no noise in the model, there are no random terms or error terms.

Also this simulation doesn't just generate residential clusters which is what I focused on, but also generates other patterns against which it may be falsified. Now they may be independent, they may not, I don't know, but we can think about that.

So move histories, how many times do people move? What's the distribution of the numbers of times that people move? Are people clustered by PP? Do more xenophobic people tend to end up in one neighbourhood, the less xenophobic in another? And so on and so on. So by specifying a complete system we're not just fitting against one output, we're fitting against all potential outputs of the system.

And finally, we're sticking our necks out and we're making unambiguously causal claims, we're saying this is how it happens, we're not saying anything is correlated with anything else, we're just saying this is how it happens. Now that means we're very likely to be wrong because of course the world is very complex, but at least we're sticking our necks out and making a nice clear claim.

Cautions – There are questions about degrees of fit, when we say two things are like and unlike, what are we saying? Simulated data. Should we be able to predict exactly where everybody lives in New York? Well no of course not. Should we be able to predict that there would be residential segregation in New York? Well that's a big weak. Where in the middle do we place ourselves? We have theories we hope of increasing quality in terms of degrees of fit.

The other thing we have to watch out for is not mistaking criticisms of the whole scientific approach for criticisms of specific methods. You might say to me well clearly people don't all make decisions in the same way, although that's actually something we tend to share with the economists in our simple models, but if each agent makes a decision in a unique way based on a unique perception of the environment then we can all just give up and go home, not just the modellers but all social scientists, we shall all become biographers and journalists.

So the important debate is not about whether there are patterns in social behaviour, but about when and to what extent different patterns exist to be found. And that debate is the one that we tend not to have, it tends to be a bit two legs good, four legs bad and I shall try and avoid that in this talk.

What about microsimulation, get back on track having provided an example. Very broadly speaking and somebody's bound to shoot me down on this, social science seems to divide into research on attributes and their relations, age, gender and so on, and research on practices and their meanings, behaviours and so on. Microsimulation I would say leans towards the attribute approach, it characterises people in terms of gender, ethnicity, age, marital status, so on and so forth. And this can be seen, I'm not just making this up, this can be seen in practices like reweighting and upweighting where you have the idea of a typical agent in a certain proportion, but also in the kinds of processes used to produce data like matching and imputation. If you don't believe that attributes and their stable relations are important then how can you justify filling in missing data on the basis of properties of existing data? If you believe for example that agents are adaptive or that there are social influences at work in these models, then it becomes much more dangerous to use matching and imputation.

To back this up, just in case I'm out of date, I used a definition of microsimulation provided by some chap called Williamson recently, which is in the International Journal of Microsimulation. It's nice to see that what I thought was true a few years ago to an extent is still true. And this also raises a worry which I'm going to discuss in the final part of the talk. I don't think it's the case that agent based modelling and microsimulation will naturally meet in the middle and I'll try and show why I don't think they will naturally meet in the middle. In a nutshell I think the reason is that behaviours as described in the agent based model that I've shown you aren't just another attribute of agents like gender or age, and I'll try and show you what I mean by that. In fact if I was wearing my woolly sociologist hat, I might say that gender isn't an attribute at all but a negotiated achievement. Now that sounds a bit horrible, what I mean by that is it's easy enough to say that particular people are male and female, obviously you just check down the underwear, but that's sort of missing the point about social causation, it's what you do with your gender that influences the social world rather than which box you sit in. So the idea that we can say well certain things are likely to happen because somebody is a woman, with the exception for example of pregnancy, is to miss the point about causal process in society.

More caveats – Beyond a certain point there's no point in trying to adjudicate definitively between different methods, it's not for me to say ABM good, microsimulation bad, what I'm going to try and do instead is say let's think about domains of application for different approaches. Most current methods don't do this, including ABM and it's something that worries me in ABM too. What I think we should be aiming at is a set of instructions on the can for our theories. This theory can safely be applied in the following empirically assessed situations, it will work badly in the following other assessed situations. I don't think we do that at all, we just take it for granted that our model can be applied and we will go with it and see what happens, and nobody ever really is going to prove us wrong. And that worries me quite a lot. We can explore the consequences of particular methods by looking at one framework from the perspective of another framework, we can go are you sure you want to assume that? The consequences of it will be this. And we also need to record constantly that each method is an article of faith. Our reasons for believing in these methods may have little to do with their truth or empirical support, they may be to do with what we learnt at university or how we see the world.

My concerns – Explanation versus prediction. Prediction is problematic in social science because pure prediction may involve no generalisation, this comes back to your point about independent support for these models, the fact that they're validated by other policies. Without some sort of explanation, without some sort of independent empirical basis for thinking that one particular flavour of model has captured the truth we can't tell, we can't tell whether this thing is just a look up table as they call it. Prediction as a goal for models also gets limited credit when there are tuneable parameters. If we succeeded in predicting well, have we just attuned the system to match the output without tapping into any underlying behaviour and how do we decide, on what grounds do we say this isn't just tuned, this has really captured some underlying truth and your answer is one possible answer, we try to find independent sources of validation.

In the simulation case what we do, the ABM case what we do, is we say OK if we're going to fit the model on some sets of properties we're going to hold back some other sets of properties in order to do our testing. So if we fit the model to generate the right kinds of clusters, we're then going to test it against its ability to generate move sequences, or its ability to identify behavioural clusters in terms of PP. Because what we're trying to do is we're trying to compare complex systems, whole sets of behaviour rather than single outputs, we're in a better position to falsify, or so the argument goes.

Concern 2 – Power and prediction. If you're building a simple statistical model, the power of your test is relatively well defined, you know how much data you need in order to decide whether or not your model fits well or not. In complex microsimulation models it's not so clear, at least from the reading I've done, how the quality of the prediction relative to the quantity of data is to be assessed. Is it very impressive that you've got a prediction of a particular quality or is it pretty much inevitable given the number of degrees of freedom you've allowed yourself in terms of noise terms, parameters that haven't yet been fitted or are only fitted in broad ranges and so on and so forth. Now this would be a problem for ABM too except that ABM doesn't try and concentrate on a relatively small number of key outputs as the test of its model. Ultimately we're always at risk of being told, well you may think your model's good but it's rubbish because I've measured it against another dimension and it doesn't hold up at all. So we're constantly vulnerable in this approach to someone coming along and measuring the simulated system in a different way and saying nah, no good. Not to say there aren't problems with that approach but it's there in principle.

Third concern about exogeneity and this is something that hasn't really come up in the presentations so far. In econometrics the idea of exogeneity variables external to the system which don't need to be modelled is an empirically determined property, you can look at the variables in the variable systems and say this can be treated as exogenous. In ABM you have to try and figure out what's exogenous by fiddling with the simulation. If you exclude from the system something which actually determines the way that it generates, you'll stop getting effective comparisons - it won't work properly. So although we can't say in advance what's going to turn out to be exogenous, if we make the wrong assumptions about exogeneity and exclude things that are actually part of the system, we'll just start getting it wrong.

At least some of the microsimulation I've looked at, appears to *assume* that certain things are exogenous. So for example if you're looking at a demographic process you just say well we won't model that. And that worries me because that, on that assumption, on the validity of that assumption hinges the likelihood that the rest of the system is going to be stable and reliable. If it turns out that benefit policies shift people's demographic behaviour and I don't know how likely that is, but it's worth thinking about, then any model which excludes that fact is not going to predict well or not going to predict as well as one that takes account of it. And that would be a nice example of the kind of discussion, we have an economic model, that's fine, there are probably major economic effects from policy change, but just suppose there were demographic effects. You know and we hear all these rubbishy stories about people trying to get on council housing lists by having babies and so on and so forth, we can't necessarily exclude those kinds of behaviours by fiat from the model.

Fourth concern – Correlation and causation. Under what circumstances should we assume, can we safely assume that missing data can be filled in on the basis of attribute patterns in existing data? It's certainly done and it's certainly a practical fix to a big problem. But can it actually be justified other than in terms of prediction which raises its own problems. If this and other things like it are done without justification, what do we do when the prediction isn't very good? We've got 5 or 6 fixes, all of which we're not entirely happy with, we don't get very good prediction, what do we do now? And this is something which also was raised in the previous talk which I think is quite important, to what extent are models calibrated in terms of individual component measurement, and to what extent are they jointly fitted? How many, at how many points does a model touch with reality? If it doesn't touch with reality at enough points then you've got a problem and in the ABM models the intention is to try and fit with reality quite closely at the micro level and then stick your neck out about the behaviour at the macro level and use that as your test to falsify against. So if you don't have independent evidence for your micro assumptions, and I raise the issue of what the independent evidence is for economic micro assumptions, you may have a problem.

Fifth concern – Noise, randomness and error. This also came up. Now clearly there are circumstances in which error occurs, people's hands slip for example, they mean to press button 'A' and they press button 'B'. But again that's quite distinct from the kind of randomness that arises from incomplete models, we just don't know how this works, so from our point of view it's random. Again if you're comparing microsimulation models and econometric models, econometrics specifies rather precisely the properties that noise or error terms must have. You can't just bung them in like blur on an

unflattering photograph, you have to say this is an error or a noise of the following kind, and I think there's a strong tendency for a certain amount of noise just to get chucked on top of a model and of the course the problem with that is if you put too much of the wrong sort of randomness in a model, can you predict pretty much anything, does it become just another fiddle factor? So how do we justify the claim that there is a particular kind of noise or randomness or error in model?

Linearity – Well even the Schelling example in its incredible simplicity and silly assumptions shows that systems can be non-linear. In these circumstances how happy are we adding up analyses of attributes? And this comes back to your point about you know we'll do this until it's proven, something better comes along. I can't get the data to prove that I can produce a better wholesale model of benefit uptake, but what I can do is show that even in these incredibly simple systems the general sort of adding up is problematic. We can't trust the social world to behave in a linear fashion, even in incredibly simple versions thereof. Now maybe like chaos, if we make the models complicated enough all this non-linearity will wash out, but I don't think anyone's very confident about that fact and I don't think we should assume it without knowing a lot more about the systems. And another way of raising my concern here about attribute based analysis, there's a whole sequence of social interactions, complicated social influence, social adaptation, people learning, if we split up the whole cloth of social interaction along attribute lines, in terms of gender and ethnicity and race and income, would we really expect the components to add back up to sensible outputs? Is that the best way of chopping up the social world and what other ways of chopping up the social world are there?

Behaviour – Economic models are not without their problems, as even economists have recognised. Does microsimulation really want to inherit by default the economic approach to understanding human behaviour? Does it want to disallow models where the agents adapt, does it want to disallow models where social influence occurs? These are pretty broad classes of models to rule out without serious consideration. There needs to be a sharper distinction made I think between what has been called accounting microsimulation and between behavioural microsimulation. In some sense accounting microsimulation is a purely technical challenge, it's a rule following system. Can behaviour be bolted on to a basically AM framework? And this is also an issue that came up in the questions, can you start with a baseline which ignores behaviour and then stick it on top like a cherry on a cake? And this is a revisit of the earlier worry whether behaviour is just another attribute like gender or class.

Drawing these concerns together – You would hope that an individual based approach clearly ought to do better than a highly aggregated one. And ABM and microsimulation agree on this clearly, but how do we make sure, using some combination of methodology and data that complex individual level models don't end up with too many degrees of freedom and end up passing the prediction test illegitimately? Now there are very bad ABM models out there which do exactly the same thing and I hold my hand up and confess it and we are working on trying to improve practice in the ABM community. But I throw this question open to the microsimulation community too, how confident are you that your model doesn't have too many fiddle factors? Could it predict almost anything if you'd tuned it right? ABM is evolving ways to handle this issue, I'm involved in a research project to do precisely this, to what extent is microsimulation also concerned about at least some of the things that I have raised?

Some constructive suggestions, just in case you're getting a bit kind of demoralised by all this. We can use agent based modelling to discover how often in some sense it is safe to use certain kinds of probabilistic models. Microsimulations can be seen as what David Hendry, the econometrician, calls reductions of agent based models. Under certain circumstances it may be perfectly reasonable to treat the birth of children as a probabilistic demographic event, but we'd like to know when that's true, we'd like to know what decision rules, what social circumstances and so on actually make that a safe assumption to make. We can do that by using ABM to generate for example transition probabilities based on certain models of behaviour. So for example in the Schelling case we can say OK let's try and come up with a microsimulation of the Schelling system. Can we generate stable transition probabilities in some aspects of this system that will allow us to understand what's going on as a first approximation? Now it may not be as good as the full description of the process, but it may be good enough for government work, as the Americans put it. But we need to check that, we can't just hope and unfortunately the few times I have done this, the news has been rather bad. It seems that even extremely simple ABMs, much simpler than real social behaviour, do not generate stable probabilities in anything much at all. Now that doesn't prove that ABM models are ever going to be better fits, but it does raise concerns about the justifications underpinning microsimulations, that there are these stable probabilities out there which can be got at. It begins to look like it might be unlikely.

Constructive suggestion number 2 – There's no reason if you want to add behaviour to microsimulations not to add agent based behaviour models, not to add social influence models, not to add adaptive agent models, you don't have to buy the economic package lock, stock and barrel unless you want to. However if one is going to do that, it's very important that you don't end up destroying the social assumptions built into these models. If you just say well we want everything to be separable, we want everything to be estimated on existing data, then basically you're ruling ABM out of court, as with any new method, if you expect a new method to improve performance on existing data, it's a hiding to nothing, the whole point about new methods is that they reveal the need for new data.

One of the main things that I think makes microsimulation much better than ABM, and this is something I'm also trying to do something about in the ABM community, is that microsimulation takes data much more seriously. There are far too many ABM models that don't even have a nod at an application domain. And I think that's something that microsimulation can definitely teach ABM and it's very important that ABMs actually start tracking real data in aggregate, even approximately, there's far too few ABM models even attempt to do that. Unfortunately this does reveal a lot that we really don't know and you know probably know the story about the drunk and the lamppost, the policeman says what are you doing here? He says I'm looking for my keys. He says when did you last have them? In that pub. Well why are you looking under this lamppost? Well there's more light here. And the design of models unfortunately tends to go to where the data is and where the data is cheap rather than to the areas that we really need to know about.

So you can see in the Schelling model that there are a whole bunch of things that we simply don't know about at the moment. But as long as ABM isn't bolted awkwardly onto microsimulation it should be possible to get it to do the kinds of things that make microsimulation useful. So my concern in a sense is that we don't end up with ABM coming in last thing and then basically destroying the logic of both methodologies, so you get bad ABM *and* bad microsimulation, it ought to be possible to get these two things to fit together properly, and I think that would be a very interesting challenge.

And finally to say this, politically you know I'm going to be no more popular saying that ABM basically subsumes microsimulation as someone from microsimulation would be saying oh we should just use agent based modelling as a kind of last minute bolt on to get a slightly better prediction.

Conclusions – The assumptions you don't realise you're making are the ones that will do you in, I hope at least some of the concerns I've raised have made you think a bit, even if you end up not agreeing with any of them. And this discussion certainly isn't meant to imply that ABM has no faults, it's got loads of them and they're not just technical, *but* that's a different talk.

Finally, Journal of Artificial Societies and Social Simulation, free on line journal of agent based modelling. simsoc, an email discussion group for social simulation where simulators of all flavours are very welcome. A textbook based on agent based modelling examples, Gilbert and Troitzcsh, many of these you may have come across already, but just in case you haven't. There's now a social simulation node as part of the ESRC National Centre for Research Methods which has an agenda of methodological improvement for agent based simulation. The example I showed is written in Netlogo which is a nice little free programme for playing with agent based models which comes with its own model library. So for any of the bits that you found interesting but weren't sure about, here are some resources.

Thank you all very much, I hope I haven't overrun by much, sorry. Questions? Comments?

QUESTIONS

Male question 1 – You're critical of models with fiddle factors that you say you know, that distort or allow a machine to fit the data, in a way I agree with you but from the agent based modelling side of course we've got considerable flexibility in how we systematise behaviour in terms of algorithms. Isn't that our fiddle factor really, that you know we don't like one decision making outlook, we'll just go and substitute another, there are thousands of different ways we might have implemented behaviour, even very simply in Schelling and we could have ended up with very radically different results.

Edmund - And this is why I think that microsimulation has a lot to offer in the sense that it pays far more attention to prediction and data than we do. And I think for example in the Schelling model people don't even get close, so you're absolutely right, I guess the issue is I think simulation is to some extent trying to put its own, ABM is trying to put its own

house in order and I wonder to extent these issues also arise in the context of microsimulation. Because people are now saying you know you've got to have some data, you've got to take account of the stylised facts in the literature, you've got to do some sensitivity analysis that shows that your result isn't just a kind of glitch on a very particular set in the parameters base. And some of these are institutional things, they won't naturally be done by individuals, they have to be done by institutions or professional bodies or whatever and say look it's just not good enough to make up a toy set of assumptions and go oh isn't that interesting? So yes you're absolutely right, but then the questions is how are these things going to be fixed? And we have some ideas about that in the ABM community.

Male question 1 – I wonder though if there isn't a danger of being too hung up on data and that danger to me is where it relates back to the previous issue in a way, where we regard the models as black boxes, the purpose of which is to fit the data and give us good predictions and all this sort of thing, and to me one of the key sort of tenets of agent based modelling is it actually matters what's inside the black box. How, you know, do you trust something that represents reality in a way you understand it and then it comes up with the right answer I suppose relative to the data that there is available or do you trust something that …

Edmund - Well this is where this issue about calibration rather than fitting comes in because the claim is that there's some sort of independent mapping of for example individual level behaviour, some insight for example into the decision making process which is attained relatively independently on the basis of which the simulation then unfolds the dynamic system producing a set of patterns which you can then test against real data. So you don't actually want a black box and the danger is if you have a black box it could just as easily be a neural net or a genetic programme or something. But the trick is how to set up a methodology that allows you to independently check your micro level assumptions. Now in agent based modelling we have *some* ideas about that, and in microsimulation under certain circumstances it does happen, you know you have to make sure that the distribution of simulated data is like real data. But I don't think either side have got very far yet with making this a kind of principled basis for model design. Some people manage it and some people don't, and it kind of you know goes along …

Male question 1 – Well I guess the closest you get are the companion model xx where they do it more in a participatory way but xx luxury of xx (34.44-48)

Edmund - Yeah and you know there's nothing wrong with predictive models as long as you don't end up with the situation where you've just got one predictive model and you're not actually comparing it with any other predictive models, so you don't really know how good it is. But you need one or the other, you either need empirical grounding or you need a prediction competition, but you can't have neither and I think you end up in the situation where you've got neither. So yeah there's something about, and actually this links into some good stuff in sociology about the meaning of explanation. James Coleman the eminent sociologist has argued that you're not really producing an explanation unless you're able to link at least 2 levels of analysis. If you just describe phenomena at a macro level you could just be matching outputs, if you only describe behaviour at the micro level then you're just doing description and journalism or sightseeing. It's only when you can link the micro to the macro explicitly that you've actually produced an explanation, and that's rather similar to the kind of methodology of agent based modelling, where you say well we'll make a set of hypotheses and the acid test to this set of hypotheses is the match to macro data, that's the aspiration. Unfortunately it doesn't often happen. We try.

Male question 2 – There's a distinction between nice to know and need to know models, a lot of the need to know models are there because people actually require some prediction about benefits or whatever. Nice to know models are things which people can do because they think it might be interesting. And in the first case you've got to work in the real world. I saw a very nice agent based model, marriage models, and I saw this paper that said we had 400 people and it was actually too slow so we cut it down to 200 and it took a week to fit this model about choosing your partners. Put them in a room with you know alcopops and a ghetto blaster you could actually do it much easier, you'd have that number. I was looking at the national population marriage models. So while it's nice to know these models will produce schedules which are realistic and produce ones which look pretty much like what we observe out there, but you have no option, you've got to use just what's out there and just plug them into the microsimulation model to produce the results. So I think, I mean sometimes they're completely different and I don't think, they're never going to match because we haven't got the computing power to use these models in real situations, so I don't see there's any sort of conflict between them, we just have to bring across whatever seems appropriate from one or the other and try to fit the best and don't somehow feel there's some sort of nirvana where we can incorporate the full agent based model within a full microsimulation model to get some kind of policy model. They're different animals basically.

Edmund - Mm, I mean you're right but I think there is a bit of a rhetoric about "real world modelling" and I don't think microsimulation models are either cheap or value free. So we say well this is quick and dirty, it gets a result in the real world, but it's like yeah but there are all sorts of assumptions built in and there are things about whether you know, whether these assumptions are actually borne out and how good the predictions really are. So I think you know if you were talking about econometric models which really are quick and dirty on an aggregate level, I would buy that, but I don't think microsimulation can say ooh we're much more in the real world, you know, we're doing this on the cheap and quickly compared to your kind of airy fairy ABM, because I just don't think it's like that. I mean big microsimulation models take years to build and I think it would be fair to say that there some questions about ex post prediction! So I think you know, I'm not sure that's a fair comparison, but I take your point, I mean there's no point in just going off into the abstract and saying well one day we might understand this. But in a sense what I've tried to do in some of this, raising this issue about you know showing that if you try and restrict agent based models to get stable probabilities you tend not to, I can't prove that a particular model is flawed because in order to do that I'd have to build an enormous model myself and there's only me, I don't have a department behind me. But I can show that if you try and restrict agent based models in terms of stable probabilities you usually don't get very far. So there is a kind of existence proof for this worry above and beyond trying to outperform specific huge models. So there are other ways of addressing this issue, rather than kind of head on competition, we can say well look there's an underlying concern here about this kind of prediction. But yeah I mean agent based modelling has suffered from being too abstract and too scientific, rather than applying (too scientific, sorry that's judgemental) too focused on the ivory tower, it would be lovely to sort this out and to play with these models, and I think it needs to get its hands dirtier, but I did say that! So I think this focus on data is quite important.

Paul: Time presses on, and I saw a couple of hands raised, but with apologies to xx, I think it would be fair to Alan to throw one final question.

Alan: No I'm following up on a few of the comments and challenges that were raised relevant to the initial conversation that we have there. You know I completely share the concern that there could be too many of what you call fiddle factors and what I would aggressive calibration! (AUDIENCE LAUGHS)

EdmundL I think I prefer that!

Alan: To shoehorn whatever subjective model that you think forms the basis of characterisation and behaviour to fit the data but you know we're hopefully responsible people, you know we would first of all you know spend as much time as possible properly and independently identifying the behavioural model we bring to the piece. And secondly having calibrated should we feel that necessary, I personally think it's critical that you then diagnose the properties of the models from the calibration process because you know the types of stochastic terms, the properties of the stochastic terms that were inherent in the model that you for example estimated. If it then comes to a calibration process that gives rise to you know completely contrary properties for those same stochastic terms then it gives you a question. Is the model doing you know not enough and is the calibration doing too much in fitting to the observed patterns of data. So actually using calibration as a diagnosis of the calibre of the model that you've bring to the piece is critical, the information content of the calibrated terms of the model gives you a form of test. So I don't think there's a difference. Aggressive calibration moves you more towards this idea of a black box, if you calibrate and you find there's very little extra information content of the stochastic terms then you would use other forms of evidence to validate the model, so it's …

Edmund - But it's quite interesting that your model does without basically all sociology and all psychology and still fits pretty well, now that could mean that all psychology and all sociology is just rubbish, but it worries me a bit …

Alan: No not at all, but I mean there are layers, I mean the idea of multi level modelling you know takes, you know individual patterns of behaviour takes you know household patterns of behaviour, we've moved to that level. If you take then localities and use that as a piece of information to you know describe patterns, I think, I suppose coming to the start of your slides I think we spent a lot of time knocking out slides which expressed concerns about microsimulation but I probably see enough for me to judge whether or not ABM can contribute to a proper adaptation of microsimulation to take account of your wider concerns. I mean it would be quite good to see something which is a little bit more involved in articulating you know different practical models …

Edmund: But to give a concrete example, this isn't, the Schelling model is a model of social influence, now my understanding of your model is that it contains no social influence at all. Now that's fine, it could be that sociology is completely wrong about the importance of social influence, but the fact that you can still get a pretty good fit and I trust you that it is a pretty good fit, without any social influence at all, does that really tell us about social influence and sociology or does that tell us about the tuning of the model?

Alan: Oh no and you know I would concede you're absolutely right, I suppose if I was to pick up in the decomposition, the variance, if I could pick up the most important influences I would probably regard what's happening you know to me as an individual in my relationship with my employer and what's happening to members of my household as maybe the prime influences on my patterns of behaviour. And then I could take it up to the next level, I think social influence is for sure something that one would like to incorporate. But I suppose if I, I wouldn't buy that a model is fundamentally challenged by the fact that it doesn't include social influence as well as household and individual patterns of influence and the influences from my employer and the influence from the state as articulated through the tax and welfare system. So you know again I have no challenge to ABM offering a contribution, I wasn't, until this discussion (!), aware that there was such a sharp divide between ABMers and microsimulators.

Talk over each other.

Paul: We'll stop the conversation there, we do have two other contributions today which I think will help to further address exactly these kinds of concerns, and the next paper is indeed one of those.

END OF RECORDING