ESRC Microsimulation Seminar Series

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

Modelling Epidemics
Wes Hinsley

[EDITED TRANSCRIPT]

Paul Williamson: Welcome to the afternoon session and three papers being presented, the first paper is being presented at very short notice by Wes Hinsley because the original presenter Neil Ferguson, is off in CDC Washington coping with the swine flu outbreak and Wes has very kindly offered to step into the breach!

Wes Hinsley: I'm going to describe to you the progress so far in this three-year project which started in January last year: the global epidemic simulator. This is an individual based approach - which will immediately strike a certain amount of terror into the heart of a simulation programmer because yes, we are going to have 6.8 billion individuals represented explicitly in this model. Quite how we'll simulate those, I'll explain. I work with a colleague, Pavlo Minayev, and we're funded by Microsoft Research, Cambridge for this project.

Our aim is to simulate the emergence and spread of an epidemic by explicitly modelling 6.8 billion people. At the start of this project we were thinking mostly about the threat of H5N1 bird flu. I guess in the last couple of weeks the world has become slightly different. But the broader aim of this platform is to model any directly transmitted disease - anything that people spread by being near to each other, influenza, smallpox, SARS; the project will provide a platform in terms of simulating what people do, so you can if you like supply the characteristics of any such disease - its incubation periods and so on.

The significant papers if you want to read up on this, are Nature Papers by Neil Ferguson and the group, where the focus has been on South East Asia, particularly Thailand and then also through the United States and the UK by air traffic. But those models were geographically small enough to fit on one computational node. The global simulator may bring this all together for example with an epidemic that starts in China and spreads to the United States.

The challenges: obviously computational performance when you've got that many agents is going to be critical. The memory usage is not just for the many individuals, but also you need matrices of how they travel which may take the form of point to point matrices which might be extremely costly – (depending on resolution). So we consider various efficiency drives to address memory cost, and reasonable performance: if it only runs in real-time then that's not really going to help any purpose on predicting what an epidemic might do and how you might stem the flow.

Data availability is a challenge: how do you build a synthetic population that is credible? People of approximately the right ages living in approximately the right places, and doing approximately the right things in terms of how they behave, how they meet each other, how many friends they've got, whether they get on planes regularly, etc. The availability of that data is extremely variable and sometimes the cost of it is quite remarkable, particularly travel data; with airline companies it appears to be of such economic value that you'd pay tens of thousands of pounds for it if you wanted it. Other places have problems where the data coverage is just poor, particularly Central Africa. We don't know much of what people do.

The complexity of the algorithm itself is a challenge, simply getting it right, making sure that we don't double-count people contacting each other in different ways. I'll show you various approaches of how people might contact each other later. We need to get the numbers right.

And then at the end, how do you know whether you've actually modelled the world correctly:- can we justify the choices that we've made? I'll hopefully get to tell you how far we are with that.

So the design of this project is to firstly create the synthetic population and work out how to represent that with a number of computers. While it will be spread in some way fairly evenly, it doesn't necessarily mean we can just chop the population into equal slices. If the computers are going to be talking to each other, which they will, then ideally you'll get

good performance if they don't have to be talking to each other all the time. By that I mean it would be good if you could put the borders between computers in the sea where not many people commonly contact each other. Alas you can't always do that:- we have to have a border running through India, we have to have a border running through China and so far, the communication cost dominates the performance of our models at those points. So some rationale for where you put your boundaries for communication between computers is important.

Having created that population, we want to avoid doing what you would call a "one-to-one", "individual-to-individual" kind of simulation where you would say for each of my individuals there is a probability that I will contact all of the other individuals. If you were to take that approach you'd have an extremely inefficient simulation. There is always that question of what such a probability would be, like you hear on the weather, "there is a probability of rain". Well of course there is, but is it zero or is it one? There is a finite probability that someone on one side of the world will contact someone on the other side, which may be very small. How do we manage that? Our approach is to overlay the world with a grid of patches, and start with a patch-to-patch probability and then to correct that later by making it more like individual-to-individual only when we need to. I'll explain how we do that later.

There are various reasons people may contact each other. It might just be a "random community contact", where people bump into each other, or there may be more detail we can add – people who share households, schools or workplaces. There's a certain group of people who will more regularly perhaps take long range travel - flights and so on.

So to begin with, to build this population we use a data set called LandScan which comes from the Oak Ridge Labs. It is essentially a matrix of (43200x20880) points, mapped to longitude and latitude. Generally it will tell you the number of people who live in patches of about a square kilometre. IThe dataset is largely sparse, so if you need to, you can store only the populated squares, reducing it from 3.5 gigabytes down to about 600 megs. It's interesting to notice as well that this [image on screen] is a representation of it plotted logarithmically; if you plot it with a linear scale on colour then you only see a couple of patches in India and China, such is the non-homogeneity of population density across the world.

So at the outset, as a computer scientist unsure of quite how to divide the world into, in this case seven different nodes, (that's the number of computers I had), I wrote a tool to do it which can make it easier, whether it's quicker in the end is anybody's guess! And here is a border that winds between Europe and Africa. It sits in the sea, so hopefully there won't be too much communication across this border, there certainly wouldn't be as much as if I sliced right across the middle where people are living both sides of the border immediately.

So then I overlay the grid of patches. The yellow ones are what I'll call local patches and this means that the node that they sit on will associate with that patch the people who live in it, whereas the blue squares I call remote – (all from the perspective of my Europe node here). And in these I just store the coordinates of the patch but not the individuals - but I do record which node has the "local patches" for those regions. So if I have an individual in a yellow patch here that wants to make a contact in a blue patch here, then I know which node to look at and I can send a request saying please find me a contact. If I zoom in a bit further:- this is the top of Egypt down here, you can just see that I've got small squares for the local node and larger squares for the remote node. All of my local patches are the same size for simplicity at the moment. My blue patches are not necessarily so; a simple divide and conquer algorithm is used (until purely blue patches remain). You can also see the sparsity - nobody lives here there are no patches. This slide is how the Africa node sees the same region – with the local and remote (yellow and blue) areas swapped.

So what you end up with when this is done for all of the nodes is: every node has a representation of the whole world but it only has the individuals for its own local portion that I assign.

We now refer back to the LandScan data set which has higher resolution than my yellow squares, in fact 20 times higher on both axes in this case. We use LandScan to position people: so within this yellow square, these little grey squares are the resolution that LandScan actually has. Within these I randomly position people by longitude and latitude.

This part may be a little more detail than we need but I'll go through it reasonably quickly. I have a kernel function which is essentially a decay function, so what it means is that people very close to each other have a higher probability of contacting each other. I use this function, along with the population density, to calculate a set of patch-to-patch probabilities.

And then I normalise this so the total of all these probabilities for any source patch adds up to 1. So what I've basically got here is a way of saying rather than doing individual-to-individual, I can go from any patch to any patch and pick one out of a probability distribution. It is however a slightly generous probability because it assumes that everybody lives right on the edges of the patches, so it's assuming that they live this distance, (Dk,k') apart. Actually they'll live further apart so we can do a ratio of these two distances, put it into the kernel function and correct the probability – called a rejection algorithm.

Plugging all of that in you can see here one of the early runs which is on one node being seeded in South America. The colours are on a sliding scale of red to green, where red means everybody in the patch is affected, green means everybody is immune.

These are the kind of typical graphs that you have in epidemiology. This is the graph of new cases in a day - newly established contacts. And this is the graph of how many people are immune and have had the flu or whatever disease it is. The textbooks have generally a much smoother curve here – there's a sort of characteristic graph that you're looking for. What it is really showing here is that the epidemic happens in two stages, it gets going and then it stalls a little and then it gets going a bit more, which looks from the shape of the geography just like it slows down as it goes through sort of Mexico, Panama and places like that where there just isn't as much space really.

So to demonstrate it further, I simulated it the other way around and seeded in Florida instead, which was even more dramatic. All that it's really proving is that people aren't uniformly spread across the planet, and if that's all that you model then you'll get very wobbly sort of graphs because their distances apart are different. So I then said well what about if we make the kernel function I mentioned earlier a lot less local - enable people to travel further. You can kind of see the spread is a little less wave like, it's got all the way around South America pretty quickly, and that gets you a much smoother curve coming out. I've enabled the people quite artificially to travel further to make their contacts. So this is more the sort of textbook graph that you see.

Tis is all one single node, so there's no computer-to-computer communication going on here and that is about the maximum that I can fit on one of the nodes that we've got. So when you get to the point of having to establish contacts on other nodes then you basically have to send requests over and receive replies back.

Perhaps I should have mentioned this earlier. There is a "magic number", R_0 which is the basic reproductive rate of a disease. It says that for every one case of a disease this is how many new cases you would expect to get. So if this value R_0 is less than 1 then a disease will naturally die out because for every infection you've got there are fewer infections coming out. In practice if it's only slightly more than 1 then eventually it may die out due to other measures coming into place or re-contacting the same person who has already had the disease and they're immune. But if it gets much higher, then you have an epidemic that's going to spread.

So depending on the value of R_0 that you put into a simulation you'll have different characteristics of the graphs that you get out. So the simulation tries to find R_0 contacts. If I'm going to find say 2 new cases then who are the most likely people I'm going to pick? When you send a request to another computer, you don't know when you send the request whether the reply is going to be positive or not, you don't know whether the contact is going to be found or not. So there's a message protocol here that basically sorts everything out and deals with the replies. There may be an extremely large number of requests - 500 megabytes is quite usual at the moment. How many people here are vaguely computer-familiar in terms of high performance code? [No-one] So I won't tell you anything much about that!

So this is what this run does: It starts in Vietnam, Asia, and is still using the same local community contact model - just people talking to each other, meeting each other fairly randomly. The simulation reaches most places but it doesn't get to Western Australia or America just because people aren't close enough to each other for that to happen in this way of modelling.

So the graphs that you get from that:- you can see it is extremely wave like, it started in Vietnam and you could see it pulsing its way through, you could see the fringes of it. This wouldn't be classed as normal behaviour for an epidemic, so it isn't convincing, particularly because no one in America got the disease.

The approaches that I could take to fix it? One is that I could use the same sort of approach with the kernel function but again make it less local, so allow people who live further apart to bump into each other more often. For example, as there are many people who live in New York the chances are one of you might have been to New York in the last little while. Has anybody? One!

Male – That's all it takes!

That's all it takes! So I could take that approach, but what you'd really want to think about flights to the USA and the people who take flights. We need more than one way of representing travel, so on the one hand to have this kernel function so people still talk to each other very locally, and on the other hand say there is a point-to-point connection between airports and countries that certain kinds of people at certain times of year will take. It's basically allowing people to move long distances.

And this is where we get into the sort of data issues. Of the data that I'm familiar with at the moment, the World Tourist Organisation has some country-to-country border data which lists the number of people from one country who go in to another just across the border, so whether they go in by plane or boat or on foot, it doesn't record separately. There are also certain data for how long people stay.

The weaknesses of it:- the USA is treated as one country, where you'd really much rather have a lot more detail. Also the general data quality is not altogether reliable, there are certain cases of inconsistency where countries appear with different values for what you would think would be very similar source statistics.

An alternative, or perhaps an additional sort of method is airport-to-airport data from ICAO – called the TFS dataset. This includes airport to airport passenger flows in some detail, but annually like WTO. There's no record of the final destination so you don't know where people's final destination was, if their flight had multiple legs.

Another thing is that when people go somewhere they don't usually stay in the airport, they travel somewhere else, often by road or by rail or they get somewhere away from the airports if they're going on holiday, so how exactly you predict where they're going to go is not trivial.

There are better data than that, but as I mentioned earlier the requests that I've made for it have tended to be very costly to the point that we haven't gone down that line. And there other places to find almost miscellaneous data; different countries and different authorities will publish different things, so there could be a huge data trawling exercise that's necessary. All this for the aim of trying to reasonably faithfully predict where people go and for how long.

So if we had that data, what we'd then do is represent people's journeys. If you know for a certain region how many people live there, (the standing stock if you like) and how many people from there are likely to travel, then you can form some simple probabilities: what's the probability if I pick someone at random that they're actually a visitor from somewhere else, as opposed to somebody who lives here? Or if I picked someone in this room, what's the probability that actually they in a certain time period will have travelled somewhere or are going to travel somewhere on a long haul flight? If you can do that, and if you can sample journey duration in some way, then you can start to build some simple travel plans like these.

For the simplest one you could say there's a probability that somebody just doesn't travel, they just live in one place and stay there. Or the person you pick could live there but they could be travelling next week, or they could travel and then come back (this is all for an infected individual, so this time frame is their infectious period.) The idea being to answer questions like who did they get their disease from and where were they? Or who are they likely to give their disease to, if they're going to travel? This probably needs timely consideration bearing in mind the last week's news of how things have spread, how we've got from Mexico to Scotland and to Hong Kong and so on.

So there are travel plans that we could build, and then if we have an origin-destination matrix, that kind of thing that could say: if you are a traveller here are the likely places that you might have gone to, or come from. Then we'd have ways of finding a contact in that location and spreading the disease that way.

So this is what I'll be doing in the next week or so, having just finished the local communication bit. The next stages will be to try and build the extendable origin destination matrix approach that says how many people go from point to point (by longer range travel). The idea is to make an extensible framework; we hadn't really contemplated Mexico in much detail until very recently and then found certain useful data on the Mexican Government website, so some way of extending the matrix arbitrarily - in the absence of all of the data that we would ideally want because the quality and availability just isn't available.

The next stage is to add extra "networks" if you like, of households, workplaces, schools - which has been done to a degree before, for UK and Thailand in previous studies, so while the previous work done on that may not have been ideal either, with various limitations, we've got a certain platform to build on.

Once the population seems to be behaving and travelling in the way that seems appropriate, we will think more about the specifics of the diseases. What I've shown you is a very primitive, vaguely flu like sort of disease, with the right kind of incubation time and the right sort of infectious period, but that's very much a guess at the set of values. So we really need to think of the particular characteristics of diseases, a lot of which is going on right now for swine flu, and that's the kind of direction that we'll be heading for in the near future.

So I think I've finished a bit early, I hope that's OK! Any questions on that? Anything that I've not covered?

QUESTIONS

Q: Just a factual question. If you go somewhere on holiday presumably you're actually rather less likely to meet the locals, on average than if you're in your own home community? I don't know, I'm just guessing.

Wes: Do you mean the locals at the destination you travel to?

Q: But the other thing I'm thinking is should you be worrying more about who else is on the plane rather than where people are going? Again I have no idea whether flu is spread more effectively in the airports than they do….

Wes: There's one particular very well documented instance, (and only one actually), of a number of people who all on one plane acquires the same disease, I can't remember how many cases there were, I think it was 20 to 30. It's certainly something that's been mentioned in one of Neil Ferguson's papers in Nature. In previous work, there was a concept of a "hotel" - very flexible kind of packages of people that are linked for some reason or another. So yes is the basic answer, that is important, and it is somewhere in the mix of networks I think that we need to model.

Q: Yeah, perhaps an idea for the future work is to simulate the effect of isolating a country - if you isolate Mexico from outer countries what will happen…

Wes: that's one of the things I should have mentioned. So the point of this simulation is to say once we've got something reasonably reliable, to then test the effect of different sorts of interventions that could be put in place. I should definitely have mentioned that! So perhaps somebody plans to travel, becomes sick at home and then still decides to travel. Or he may on his own change his plans and decide not to travel after all. Or an external intervention might happen that - if it is recognised as an epidemic then the governments may say we're not going to let anybody travel, or they may close schools or put a curfew on people; various measures could be put into place to try and stem the flow.

So this I suppose is really what the project will be for in the end – to give a way of assessing what interventions would be most effective, and with what confidence, because as you can tell it's quite an inexact sort of model currently, because of the data problems and because we're trying to model individuals realistically, but at the same time accepting that you can't exactly model what every person will do. So it's always with certain measures of confidence and probability. But to be able to provide recommendations with probabilities and confidence of what we we're to do to try and stop the epidemic spreading: that's what we're after.

Q: Have you been able to access a contact tracing data, maybe it doesn't apply to the kind of diseases you're looking at.

Wes: Contact tracing?

Q: You know if somebody's seriously ill they actually force people to sit down and make a list of all the places they've been, in some cases they will treat people that they've had close contact with, they used to do it for STDs, obviously then you're sleeping with them, but I wonder if any of the diseases you're looking at: they might have instituted contact tracing in which case you have a very small number of very detailed bits of data about how far apart people were, who they were, which might allow you to sharpen your kernel a bit

Wes: We could do, there are lots of people in the department who work with more STDs who probably had …

Q: But I wonder if they did this with SARS for example when they were initially worrying about it, when someone got sick, did they say well where have you been for the last 3 weeks? Can you name various people that you've been in loads and loads of contact with then it might... It's not going to be a large quantity of data but it might be rather useful.

Wes: Yes there are other tracing sources of data too - some mobile phone networks have data on where people have travelled to and when they turned their phone on. Again I'm not sure how much access we've got and at what cost.

Q: I have 2 questions, first thinking about computer efficiency, have you tried out how much you could decrease the number of agents whilst getting very similar results. And the second question is have you tried comparing these against some real historic epidemics?

Wes: Right, so the first question, have we have tried with few agents? The answer is at this stage the results I've shown are actually a little bit of a cheat because I've only got 7 nodes and I'd really like more, so we already run with a scaled down population, with fewer agents than I'd like to.

Q: By the way how long does it take for a simulation?

Wes: This took 18 hours running on 7 64 gigabyte nodes. I don't know as yet what difference reducing the agents makes to the results. At the stage that I'm at with epidemiology in general, I don't have enough familiarity with what we're really looking for yet, so it is a good question. I don't know at this stage whether you can get away with scaling down the population, but intuitively I'd probably say that if you can model it, if you've got enough resources then you probably should philosophically use all the data (and individuals) as close to reality as you can. How much of a difference it makes – I don't have a feel for that as yet. What was the second question?

Q: Validation.

Wes: Validation. With this work, not yet. But there's a lot of validation that goes on in the epidemiology field to epidemics that were very well documented, particularly one in 1918, the influenza outbreak where in certain areas the paperwork was maintained to an extremely high quality. So regarding the general characteristics of flu, there's still ongoing work, there are still new papers reanalysing that data and working out the value of the basic reproductive rate, the time periods and so on. So I suppose by the nature of it there's only a limited amount of validation that we'll be able to do with this, because when it's in use it will be taking the generic stuff we know about flu, putting in the latest data that we've got, the latest figures and then really almost on the spur of the moment trying to create some useful predictions, some useful guidelines. Whether what's happening at the moment with swine flu will create us some data to validate against I'm not sure, it might well do. I should think there's a lot more focus today on trying to keep data and track down the spread because with SARS and with the bird flu risk too, there's more awareness now that it's something that we may seriously need to have access to at a key moment.

Q: Even with quite coarse data you could get quite interesting validations - could you reproduce the country sequence of first recorded infections or could you reproduce the rank order of case numbers by country, because that's much more minimal data than actually knowing exactly who has got it and where, but it's still quite a tough test of: does this stuff just spreads out country to country in distance order, or are you getting cases where you've got no cases at all in country A and then suddenly it jumps a country and then it shows up in country B. So you've got quite a strong falsification even with quite minimal notification data.

Wes: Yes, so it would probably be along the lines of seeing if the general behaviour of simulations that come out of the model are able to do. As random numbers are involved in the simulation [and ensembles of runs therefore] it would be unrealistic to expect you could exactly mimic what's happened, but to get some of those characteristics …

Q: Because that seems to be the validation you do get, 4 cases in Scotland, 137 in Mexico, 10 in Belgium, in a particular order, people are unfortunately unfolding in that sort of order.

Q: Have you considered local transport as a context for contacts? I'm thinking of cities where there's a high proportion of public transport travel, as in very crowded conditions, mentioning no names, as compared with those in large cities where much more travel is in private cars. I would have thought that for many people that there would be far more contact with potentially infected persons while travelling to and from work or school, than while they're actually there. Supposing that many people are not working in public offices, but in fairly close communities.

Wes: Not specifically [it is represented somewhat in the kernel function] - but we could do more detailed work in due course, [subject to data availability]. The tube?

Q: For example, just thinking of like New York and large cities. For a fairly wide range of cities there is information available on sort of overall mode split and what the proportion of people are using cars or public transport.

Wes: Yes, it's a good point.

Q: I think it's interesting talking about that model because it looks to me well placed to model a very large scale epidemic and it seems to be saying that essentially everybody gets it at some stage. But in the early stages, something like swine flu you could easily see that part of the countries it's in, completely different if half a dozen countries and you've got a case of it ending up in a particular country, the chance of it actually breaking out would be different. So if you sort of repeated these simulations, if you have sort of longer distance transmission you'd get completely different results in each nation and I just wondered about the presentation of those, what would you actually present? Would you sort of say I'll do some distribution of things or would you say, I mean I just wondered: when everyone gets it, but it's much easier to present the results. What's the communication policy about this?

Wes: Right, I don't know too much about the communication policy itself but in terms of getting the results I think it would always be running quite a number of instances and getting a spread of what would, what could happen from a particular infection point. So I think we'd always be talking about probabilities of where further infections could happen. In terms of the communication further than that, I don't think I've got any knowledge on that!

Q: Would they take a distribution of people who live in Mexico where they go and say that's the distribution of whether the infections may break out for example.

Wes: I would think so for a start. I suppose by the time you get that much data you may have further information already, so I suppose it's more for the kind of thing, if we may be able to provide in due course, answer questions like what should you do in Mexico for instance for a start, to try and address what's happening there? So especially for countries that don't have their own modelling facilities and may not be thinking about themselves, I suppose the question is along the lines of should schools in Mexico be closed? Should public transport networks there be closed? And the movie is in a way the sort of sales pitch and the sort of way of convincing people that this is what will happen is nothing gets done. But at the finer level I suppose we're looking at, with it being global, the best available data, what could you do in any of these places, wherever it breaks out? What would be the best policies to employ there? Does that answer it roughly? Roughly! Good!

Q: I mean in terms of that, in terms of whether you close schools or not, it looks like you know on what you've got, it doesn't matter what you do to protect it, the world's going to get it so how do you … So practically what difference does it make?

Q: You haven't tried any policies have you?

Wes: No.

Q: If 3 weeks ago Mexico closed down for a week, if you can simulate that, if they closed down for a week then everybody wouldn't get it… (talking together – general point: If R_0 is greater than one, the infection spreads - but policies can change the effective value of R_0)

Wes: If I infect everybody in this room but we lock the door and then we deal with all of our illness! And then we go out well and healthy, then we will have changed what would have happened. So I suppose it's just another way of saying that there are patches on the movies that are not infected and haven't been reached because of their locality, although there are some people living there. So if there number of infections, if you can change the sort of incoming assault of virus coming into place then yes you can …

Q: Well it would be interesting, when you get to the point where you can start putting in the counteractions in the model that would be really interesting to see. I have one other question. You said that was, this was community contacts only, and that's just you've got a probability of contacting somebody else and the probability changes depending on the distance, is that right?

Wes: You sample for a distribution of how many contacts one infected person is going to make and then you probabilistically choose from patches which are the likely ones. So in a slightly monte-carlo like way you pick them and either accept or reject that contact, depending on probability until you have enough contacts to satisfy what …

Q: So how does it get to New Zealand but it doesn't get across to America or the West Coast of Australia?

Wes: Partly probability based, so partly there is a very small probability that somebody could have jumped this gap locally, there's nothing in the model that prevents that except the fact that the probability is low.

Q: I was going to ask whether your algorithm actually identifies that 180 degrees west and 180 degrees east are the same thing.

Wes: Yes it does!

Q: They do? I'm surprised it gets to New Zealand but it doesn't in fact get to Alaska.

Q: Well there's no people in that corner is there, that is sparsely populated.

Q: So I mean that's behaving as if there is air travel around, so I just wondered if there's something way out of the problem about not knowing about the data.

Wes: No I think it's just down to chance that this has managed to, on this instance …

Q: They probably don't show up at that resolution…

Wes: I suppose the misleading thing as well is that, if I go back to that Land Scan image - the population density of Australia is generally extremely low, so there aren't really any major pulling points other than down the sort of Sydney region. So it's a case of the distance from there probably to Alaska, bearing in mind the low population density – that's probably what does it. But on a different instance, if I were to run it a number of times eventually I would probably get an epidemic that did seed and make the gap to Australia, if you run it enough times anything may happen probabilistically.

Q: So when it comes to doing, testing out the policies or whatever, do you have the capability to run the second run with the same random numbers that you do the first time to see if the impacts are just your new random numbers or if it's your intervention?

Wes: That's a good question, I think there probably should be - there should be a way of, but I think at the moment to get it performing well, maybe that doesn't work right now, but yes there should be a way of comparing it with the interventions and without, in a reliable, fair way so you know it's the interventions that make a difference.

8

Q: Can I make an observation following on from that. In our model we found it very difficult in dynamic microsimulation to actually ensure that you did put the same random numbers in all the simulations because the situation where you have, because you're running a different simulation and your objective equations are different, what tended to happen was in one scenario a person died because we changed the mortality rate, therefore we didn't need to use a random number, you could use for the labour market participation, and somebody else got that random number Very difficult to make sure you use the same random numbers, the same events occur.

Wes: Yes.

[END OF RECORDING]