



Public Administration Select Committee

Oral evidence: [The Future of the Census](#), HC 1090

Tuesday 25 February 2014

Ordered by the House of Commons to be published on

Tuesday 25 February 2014

[Watch the meeting](#)

Members present: [Mr Bernard Jenkin](#) (Chair); [Alan Cairns](#), [Paul Flynn](#), [Kelvin Hopkins](#), [Greg Mulholland](#) and [Mr Andrew Turner](#)

Questions 1-134

Witnesses: **Professor Jane Falkingham**, Director, ESRC Centre for Population Change, University of Southampton, **Professor Chris Skinner**, Professor of Statistics, LSE and **John Pulliger**, President, Royal Statistical Society, **Keith Dugmore** Director of Demographics User Group, Demographic Decisions Ltd, **Sarah Henry**, Head of Intelligence and Performance, Manchester City Council, **Juliet Whitworth**, Research and Information Manager, LGA, **Piers Elias**, Demography and Modelling Officer, Tees Valley Unlimited and **Professor Les Mayhew**, Professor of Statistics, City University, gave evidence.

Q1 Chair: Welcome to this evidence session on the future of the census. Could I ask each of you to identify yourselves for the record?

Professor Falkingham: Good morning. My name is Professor Jane Falkingham. I am the Director of the Economic and Social Research Council-funded Centre for Population Change at the University of Southampton. I am also a Professor of Demography and International Social Policy.

John Pullinger: I am John Pullinger. For this purpose, I am acting as the President of the Royal Statistical Society.

Professor Skinner: I am Chris Skinner, Professor of Statistics at the London School of Economics.

Chair: I am very grateful to you all for coming today. It would be very helpful to us if you could keep your answers short and crisp, and for colleagues, we will endeavour to do the same with our questions. If things are dragging on a bit, I will pull you up and press on. Can I also mention that if you are following this session on Twitter, there is a hashtag we are using: #census? I would like Mr Hopkins to open the questioning on the value of the census.

Q2 Kelvin Hopkins: It has become fashionable recently to bash the census. I do not adhere to that fashion myself, I have to say. Is the census good value for money?

Professor Falkingham: Yes, I believe so. The last census came out at around 80 pence a person over the 10 years, so it cost about £8 a person. In terms of the uses that the census has been put to, there has been some very interesting research by the Market Research Society, where they have talked to their members about how the census is used. They estimate that it probably brings a net benefit of what I think was around £300 million a year to British businesses, local authorities, Government, etc, actually using the census to better target services, and to better target where businesses are located.

Q3 Kelvin Hopkins: So it is wise not just to look at the cost of the census, but actually the value in money terms to all the other organisations that use the census?

Professor Falkingham: Absolutely.

Q4 Kelvin Hopkins: There is another question. I am seriously concerned about the loss of time series that go back for well over 100 years. It is of interest to academics; it is about knowledge about our society and trends in our society. Would that not all be lost if we lost the census?

Professor Falkingham: I think so, yes. A lot of the work we do in the Centre for Population Change is trying to locate current changes in the UK population, and locate them within the historical trend, but also then to think about where we are going in the future. In order to have better population projections about where we are going, we need to look back into the past. We use a lot of the time series data from the census to inform our projections about where we are going in the future.

Q5 Kelvin Hopkins: In terms of the costs, you mentioned a figure. Another figure is that the cost of the census is less than £1 per person per year, which is a tiny fraction of other public expenditures. Do you know how we compare with other countries? I understand that we are actually relatively efficient compared with other countries.

Professor Falkingham: We are amongst the more efficient. It is more expensive, for example, in the United States than it is in the UK. It is difficult because a lot of our European neighbours have moved to a slightly different way of doing the census. They have the benefit of population registers. They know people are obliged to register when they move home, etc. so they have complete registration of the population. Of course, there it is a bit cheaper.

Q6 Chair: Could you just explain the difference between a population register and an address register?

Professor Falkingham: If we had compulsion in this country, where you actually had to register every time you changed addresses, an address register and a population register would be fairly close together. However, in the UK we do not have that. An address register is just a register of all the addresses in the country, but we do not necessarily know who actually is living at those addresses.

Q7 Chair: Mr Pullinger, do you want to answer the question that has already been asked?

John Pullinger: Professor Falkingham has given a pretty clear answer with some numbers around it, but I will just supplement it. We are using census numbers to allocate tens of billions of pounds of money across the country. We are using it to make investment decisions for billions of pounds of private sector investment, such as where to locate shops and businesses, as well as other infrastructure such as schools and hospitals and so on. We are using it to calculate rates. With things like teenage pregnancy, for example, you can count the number of girls presenting as being pregnant, but unless you know the number of girls in that age group you cannot effectively calculate a percentage. The census is the bedrock for that. You are using it to test what sort of government interventions work in different areas. If you have done something in Manchester and something else in Liverpool, the census data enables you to compare the outcomes in both of those places.

The final thing I would mention is that so many of the statistics that we use on an everyday basis are based on surveys. For example, the labour market data came out the other day and tells us that we have record numbers of people in employment. You can only gross up from the survey numbers to the totals in the population by knowing what the population is in the first place. All of the numbers that are in the statistical system in one way or another tend to depend on having this bit of infrastructure in the middle of it. That is the census.

Q8 Kelvin Hopkins: I have one more question, and really it is a positive question about the census. Other forms of measurement might give a particular, interesting view of something, but the great thing about the census is it gives longer trends over time. These are really valuable and really interesting in my view. Would you agree?

John Pullinger: Yes.

Q9 Chair: It is reported that the value of the census is worth far more than it costs in terms of the public benefit of this data, and indeed the data that the private sector uses. Why do we not set the census up as a profit centre, so that it charges other Government Departments for the use of this data or for access to this data so that we can demonstrate how valuable it is? What have I not understood?

John Pullinger: If there was an effective market mechanism that would enable that to happen, then you would make some money out of running the census.

Q10 Chair: Yes, but if Government Departments believe this data to be very valuable, why does the ONS not charge other Government Departments for it in order to demonstrate this? It would be an internal accounting mechanism, but it would demonstrate the value to Government Departments of this data. Or perhaps they would not buy it, because they do not really want it that much.

John Pullinger: I can only go on the experience that I have had in another life when I worked at the ONS. There was a period when, for many of the surveys, exactly that philosophy was promoted. It worked quite successfully if there was a single Government Department wishing to buy a particular survey. However, when you wanted to do

something that required a number of Government Departments to do it, it was very difficult to set up a system, with some people who were free-riders on it if the others came into it, that would fairly reflect the value that each organisation got from it.

The one that caused me the most grief was the Wealth and Assets Survey, which is now established. The first time we did it, though, there was no central money for the ONS to do it, so we had to do it by passing the hat around Government Departments. Every Government Department said, “We have no information. We have information on income, but we have no information on wealth, and we need it”, but not one of them would put their hands in their pockets. We went round for years and years, and it was just an exhausting and rather expensive exercise. In a system where there is hard charging for all sorts of services, it might be possible. I just do not think we are in that system yet.

Q11 Kelvin Hopkins: Supplementary to that, is it not part of our democracy that the citizen has access to this information? The great thing about the census is that it is a public document that we all can see and read about to know about our society. Restricting it and selling it would make a real difference to our democracy and access to this knowledge.

John Pullinger: If I may, I do not think that is what the Chairman is suggesting, is it? You are not suggesting it would be sold; you are suggesting that if it is valued in Government Departments, they ought to be able to pay for it. If I may, I will respond to your question as a citizen now. In my private life, I am chairman of a small company in my town and we were threatened by a very large development right next door to where we are. It was only by using census data that I was able to put together the information I could take to the council’s planning committee and say, “Look, this development is not consistent with your own policies. It would not be good for the town and it will not be successful”. The data enabled me to have a voice, which meant that the building that is going on at the moment is a much more suitable development that I think will be better for the town, and certainly better for the organisation that I run. As a citizen, that is a real public service to me, certainly worth a pound a year.

Chair: It is only the census that provides that local, granular data.

John Pullinger: Precisely.

Q12 Chair: I have got just one other question arising from this part of our brief. The Treasury Committee called for a national address register as long ago as 2007. Do we now have a national address register?

John Pullinger: I do not think we have a national address register that is going to work yet for this purpose. ONS spent £12 million in the last census, only two and a half years ago, to create one that would be successful for the census, so certainly in 2011 there was not one. There are a number of developments in place that are moving us towards it. The GeoPlace system, which brings together Ordnance Survey with the Land and Property Gazetteer, seems to me from my understanding to be a promising development.

On the other hand, the privatisation of the Royal Mail, which owns the Postcode Address File, potentially means that this will be less accessible to the public and to the wider community now than it might have been in the past. The thing I would wish to stress, and the thing the Royal Statistical Society stressed in its response to the ONS consultation, was

that the ONS needs to have access to a properly open address register of all the addresses in the country.

Q13 Chair: How hard would it be to create that data file?

John Pullinger: Many people have argued for it for a very long time—15 years to my own knowledge. It has foundered on the fact that there are a number of different organisations with responsibilities there.

Q14 Chair: How do you run a traditional census without an address file?

John Pullinger: You cannot.

Q15 Chair: We have to create one every time we do a census?

John Pullinger: That is what has happened up until now.

Q16 Chair: That is part of the expense?

John Pullinger: It was last time, yes.

Q17 Chair: We would need an address file anyway?

John Pullinger: Yes.

Q18 Chair: So in fact, the cost of the census is overstated because part of the ancillary benefit of running the census every 10 years is that we create an address file?

John Pullinger: They only create it for that day.

Q19 Chair: They only create it for that day and then it is lost again?

John Pullinger: Yes.

Q20 Chair: That seems to be extraordinarily wasteful. Do we agree with that?

John Pullinger: Yes.

Chair: Professor Skinner, you are nodding.

Professor Skinner: Yes. I think it is critical. It was critical for the 2011 census, and going ahead; so yes, it is.

Q21 Chair: How much of the cost of the 10-year census is creating the address file?

John Pullinger: I am not absolutely certain, but my recollection is it was £12 million.

Chair: We can move on to the next section. Mr Flynn will be asking about timeliness and reliability.

Q22 Paul Flynn: Professor Falkingham, you say that there is £300 million benefit to the public, but it cost the public £482 million for the last census. How did you explain this £300 million benefit? Who claimed it and on what basis?

Professor Falkingham: No, that was £300 million per annum benefit, so if you gross that up, that is £3 billion.

Q23 Paul Flynn: Who made that calculation?

Professor Falkingham: Those were calculations made by the Market Research Society.

Q24 Paul Flynn: Who was paying them to do this?

Professor Falkingham: This was part of their submission of evidence to the Office for National Statistics. They wrote to their members. Many of their members use census information for making decisions, such as Mr Pullinger said, around locating supermarkets, etc.

Q25 Paul Flynn: Thank you, I appreciate that. Mr Pullinger, what percentage of a sample would you require normally to reach a reasonable conclusion when you were down at the ONS? In how many cases did you have 100% interviewed on a sample? If you were making a report about book makers or candlestick makers, you would take a sample of the proportion.

John Pullinger: Yes.

Paul Flynn: What would be regarded as a reasonable sample?

John Pullinger: It depends what you are trying to measure.

Q26 Paul Flynn: Would it be 10%, 1% or 0.001%?

John Pullinger: If you want to have some degree of accuracy, you need to have enough people in your sample.

Q27 Paul Flynn: Yes, but what percentage? We understand you need accuracy, but what would be a reasonable percentage that you work on routinely in the ONS? I would suggest less than 1%.

John Pullinger: If you are trying to get a survey of the general population and an estimate of a characteristic it would be very small.

Q28 Paul Flynn: Yes, but you are not answering my question. In your work in the ONS, what do you usually work on and say, “Well, that will give us a reasonably reliable picture”? Would it be 1%, 0.1%, 0.01%?

John Pullinger: For the general population, it could be 0.1%. 5,000 out of—

Q29 Paul Flynn: Is there any other area in which you look for a conclusion based on 100% of the sample except the census? There is not; I will take that silence to mean that there is not.

John Pullinger: I cannot think of any, no.

Q30 Paul Flynn: We know there is not. In biblical times they went round, because that was the only way they knew how to do it, to count the population. Now we know, if you take a representative sample, you will get the same result as if you took 100% of the sample. There would be very little difference in it. Is that not true?

Professor Falkingham: It depends on what you are looking at. For example, if we are interested in looking at—

Paul Flynn: Well, all right. If you—

Chair: Mr Flynn—

Paul Flynn: We are not getting answers.

Chair: Let Professor Falkingham answer.

Professor Falkingham: It depends what you are interested in. If you are interested in looking at individual groups like ethnicity or lone motherhood, then you actually need a bigger sample because they are smaller—

Q31 Paul Flynn: Okay. How many pregnant women would necessarily declare they were pregnant, Mr Pullinger? This was one of the examples you gave. Is there not a degree of distortion likely to emerge from that from those who are concealing their pregnancy, to use your example?

John Pullinger: No. The example I cited of teenage pregnancy was to get a rate. You know their number—

Paul Flynn: Which is likely to be unreliable.

Chair: Please let Mr Pullinger answer.

John Pullinger: You can count the number of girls who appear in a hospital as pregnant. That is relatively straightforward and you can get 100% of those from the hospital records. In order to calculate the percentage of all girls that become pregnant at age 16, for example, you need to have a good, robust estimate of the total number of girls in that age group. That was all I was saying in that particular example.

Q32 Paul Flynn: Do any of you believe there are 390,000 Jedi Knights in the country? I take this silence to mean, no you do not. This is an example of manipulation of the census, would you not agree?

John Pullinger: The census in 2001 asked people to self-report what they considered to be their religion.

Q33 Paul Flynn: 390,000 said “Jedi Knights”. Do you believe that? Do any of you believe it?

John Pullinger: I believe that 390,000 people reported—

Q34 Paul Flynn: No, that is not my question. I am asking you whether you believe their religion is that they are Jedi Knights. They are not non-conformists, Plymouth Brethren or Parsis; they are Jedi Knights. We know what happened. Do you suggest that the census, instead of being a reliable form of information, as it misses out 3 million people who refused to take part, is an unreliable one, open to manipulation by groups like the Jedi Knights?

Professor Skinner: There are separate issues there. There is the question about the representativeness of the people who are included. Take the issue of sampling where you have a sample of the population rather than 100% and you are looking at very small groups and, in particular, very small areas. Let's say the sample within those is 4% of the population. The ONS calculates that if you were going down to an area smaller than 800 people, the number of people you would have there would be so small that you would get unreliable estimates for sampling variation. On the other hand, there are measurement issues that the questions that go into the census, or indeed any other survey, can suffer from people—

Q35 Paul Flynn: Would you agree that 3 million people, representing about 5% of the population, probably have a lot in common? They come from similar backgrounds, possibly. Then we have got a very large distortion of the census because people might give information to other people, but they are reluctant to give it to the census. Is that not true?

John Pullinger: It would be true if that is what the census showed, but the census in 2001 and the census in 2011 made estimates for the ones who were missed. Very shortly after the census was conducted, a further survey was done to check out the people who were missed in the original census. An estimate was given down to the detailed characteristic level.

Q36 Paul Flynn: So this £0.5 billion conclusion is based partly on an estimate. What do you think would be the difference in the results if the census sought to interview 0.01% of the population?

Professor Falkingham: If that was the case then we would not be able to have local statistics. They would be too unreliable. The confidence intervals would be too great. If you wanted to have information for your parish council, it would be impossible.

Q37 Paul Flynn: Political parties spend a great deal of money following what opinion polls tell them. Those opinion polls are based on a tiny percentage compared with the census. Big decisions are taken on the basis of 2,000 or 3,000 people at the most. Who is wrong? Is it that the census, which has a huge sample, is unnecessary and hugely expensive? Is it not true that statistics tell us that if you get a representative sample, you get the same result as if you interviewed the entire population?

Professor Falkingham: No. I think you are talking about two different things.

Q38 Paul Flynn: We are talking about wasting £1 billion of public money in the next census, so I think it is important to know whether it is necessary or not. Do we get £1 billion worth of improvement in the figures?

Professor Skinner: The majority of government statistics is based upon sample surveys, where everyone is producing estimates at the national level or large regions, where it is true that one can take very small percentages. It is an achievement of sampling methodology that one can get very accurate estimates in that way, but the census is the exception.

Q39 Paul Flynn: Can you answer the question that Mr Pullinger refused to answer? What is regarded as a reasonable sample to take in order to get a reasonably accurate picture? Is it 0.1% or 0.01%? What do we normally do? What does ONS do on a routine basis to sample?

Professor Skinner: So one calculates some level of accuracy. Professor Falkingham has mentioned the confidence interval.

Paul Flynn: What percentage would you regard as being—

Q40 Chair: Can I interject at this moment? My understanding about survey data is that it is not the share of the population that influences the accuracy of the survey; it is merely the numbers. To have a reasonably accurate survey of any group, you need at least 1,000 to be surveyed of that number—I think it is 1,017 or some curious statistical number—to get a 95% confidence interval. Unless you have interviewed 1,017 pregnant teenage Muslims in the London Borough of Hackney, you have not got an accurate sample. Am I correct? Have I understood that correctly?

Professor Skinner: Yes, broadly.

John Pullinger: That is broadly correct.

Q41 Chair: So just interviewing 0.1% of the population would be good for a general picture of the population, but how good is that for telling us how to provide housing in Hackney for single mothers?

Professor Falkingham: Absolutely right.

Q42 Paul Flynn: What is concerning is that Westminster City Council complained that there was a huge disparity between the census figures and the figures that they as a local authority have. That is another weakness, surely. Why do people have to wait two years for some information in the census?

Chair: There are two questions there. I am interested in both those questions.

Paul Flynn: I am hoping by asking a lot of questions that one or two might be answered.

Chair: The first question was about Westminster. Here we have a census that Westminster vociferously complained had undercounted their population. What is your answer to that?

John Pullinger: Some places are inevitably much more difficult to collect information from than others. Westminster has the most complex and difficult population makeup that you could image.

Paul Flynn: It is very untidy, yes.

John Pullinger: My understanding is that in Westminster, the ONS has tried to work with the local council to ensure that the estimate for Westminster's total population uses all the information that it has at its disposal. The census never claims to be perfect. You have mentioned the missing people in it, but what ONS—

Q43 Paul Flynn: Is it a fair description to say that £0.5 billion would be spent—£1 billion next time—on a census that often reports late, is distorted by people with special interests, does not give an accurate picture in many other ways, and is in fact deeply unscientific because it takes a huge sample, which we know in modern statistics is not necessary, and hugely wasteful?

John Pullinger: No, I do not think that is correct at all. The job that the ONS has is to create what is effectively a piece of national data infrastructure to try to get, on a consistent basis, across the whole of the country and to a similar level of quality, information about a vast range of characteristics, which Government Departments, community organisations, businesses and various others, say is necessary to make important decisions. That is not to say there are not all kinds of other surveys that are needed, such as opinion polls, but the census provides bedrock on which so many of those other statistical calculations are based.

Q44 Paul Flynn: It is crumbling bedrock with great sinkholes in it, I think.

John Pullinger: It is bedrock that is becoming harder to keep solid. The work that the ONS has done in this preparation for the consultation they are now doing is recognising that it is becoming harder to do a census purely in a traditional way. Sometimes it is going to be necessary to think laterally about whether there are cheaper, certainly, but also more creative ways of counting the population, particularly in complex areas such as Westminster, in a timely way. To answer your second question, the census does come out much later than would be desirable. If it is possible to harness administrative data that is available in real time then so much the better. I do not see the ONS resisting that.

Q45 Paul Flynn: Would you confess that you are all part of a statistical establishment who love what is going on? You feel inert about changes, I am sure. Would you not just agree that if we had a continuous census, taking place every six months or every year, on a tiny sample compared to the total sample, we would get accurate, timely, reliable figures, at a fraction of the cost of the £1 billion for the next census?

Chair: That is a yes or no.

Professor Falkingham: You would get some information, but you would not get the information at the level of granularity that we need to make decisions around allocation of public expenditure.

Paul Flynn: Or the number of Jedi Knights. You would not know that, either.

Chair: Mr Pullinger, yes or no?

John Pullinger: No.

Professor Skinner: No.

Q46 Greg Mulholland: To follow up on Mr Flynn's question on the accuracy, is it not the case, in a very strange reality, with the huge advances in technology over the 10 to 20 years between the most recent censuses, that actually they are getting less, not more accurate? This is at a time when it should be easier to collect data and, indeed, many people are more used to providing data. In 2008, when the Treasury Committee looked at this, Professor David Coleman said that the UK had not had a satisfactory census since 1981. He said, "Both the censuses of 1991 and 2001 deviated considerably from the corresponding population estimate and, despite every effort, each has turned out to be in error, or at least remain controversial". That has potentially been the case even more with the 2011 census. Can we expect the 2021 census to be even less accurate despite being much more expensive?

Professor Skinner: Around the world in Western countries there has been a relentless decline in public co-operation in both surveys and censuses. In 1991 and 2001 the big issue to do with the accuracy of the census has been the undercount and the implications of that for the estimates. That became significant in 1991 and very noteworthy in 2001. If one looks at surveys between 2001 and 2011, one saw the response rate from the Labour Force Survey, for example, decline from 70% to 50%. Against that backdrop, 2011 was remarkably successful. It managed to maintain the response rate, estimated at 94%, that was achieved in 2001, whilst at the same time reducing the variation between parts of the country. That is the critical thing that can be damaging in terms of this undercount: if one were finding some places are undercounting much more, the various uses of the census could be damaged.

From my feeling, the 2011 census was pretty successful in resisting some of those trends. A lot of work has been done into how one can measure the accuracy. In my broad view, I do not think that the census is broken or in dire straits, in terms of the model. Nevertheless, these are clearly challenges. If one sees the developments around the world, there are expectations that in 2021 it could be harder still to get that co-operation. It does seem to make sense to think about alternative options.

John Pullinger: We ought to do much better. The Treasury Committee was right; we have got data coming out of our ears in all sorts of places, and we ought to be able to bring them together to create the kind of data that we need from the census. That is tantalisingly close, but it is not there yet and that is the risk. It would be reckless for us to get rid of the census until we have demonstrated that the kind of information for which we rely on the census can be provided reliably by another means.

We spoke earlier about registers. Without an accurate register, it is very difficult to be confident that we would know where everybody was. Without legislation that enabled ONS to access the various administrative data that you would need to put together the data from other sources, which countries, such as Scandinavian countries that Professor Falkingham mentioned, have had for decades, we could not just turn that on. Even if you

legislated for it, there is a big question about public acceptability: whether people would want the data from the GP register to be merged with data from the benefits system and the school system to create a kind of database from the census. Often, people do not like the census when they fill it in. As Professor Skinner says, there has been a declining acceptance of it. The risk is that people would be even more anxious if a central public authority was bringing together their data from all these other sources without their consent. At least they have the opportunity to fill in the census form.

From a statistical establishment point of view maybe, I would hope that that public argument would be able to be won. I do not think it has been yet and, until it is, there is a real risk that if we got rid of the census, despite the fact that all this other information is out there, we would be unable to bring it together in a form that would meet the requirements that we have all mentioned earlier in this discussion.

Professor Falkingham: To add to that, the quote from Professor Coleman is absolutely right. There were issues with the 1991 census and the 2001 census, but we learnt a lot from those. As Professor Skinner said, there was a lot of work done in 2011 to actually target the areas of the country where we knew the undercount was greater. In fact, there was a lot of work done with local authorities. What has been interesting is that the local authorities are working in tandem with the Office for National Statistics. For example, Westminster City Council has been making available some of its administrative data in order to work with the ONS to improve the estimates of the undercount. We can draw strength from some of the administrative data to improve the census estimates. Equally, we can draw strength from the census to improve some of our administrative data.

Q47 Mr Turner: In 1941, there was no census. What was the consequence of that?

Paul Flynn: The war went on for another four years.

John Pullinger: The difference was, after the war we had national registration. We had the information collected through an administrative means. We had a 20-year gap in the time series and Mr Hopkins is right that the time series is very important. However, the Government invested very heavily in data collection immediately after the war through the national registration system.

Q48 Chair: Moving on to this question that you have already touched upon about the future of the census, what do we think of the consultation produced by the ONS and the two options being offered?

Professor Skinner: Perhaps I can start. I led a review that was commissioned, in fact, by the Royal Statistical Society, of the methodology of the work that the ONS had been doing leading up to those options, and the way it set out those options. It was just on issues to do with reliability and accuracy; it was not to do with any attempt to weigh up user benefits.

I can just summarise roughly what we concluded in this review, and what I personally think are the implications for the future. We felt that the work that the ONS had done in identifying different options and coming down to these two options was reasonable. We did not have any major objections to that. We saw one option that I think is sometimes called the online census option as a pretty natural development of existing census methodology. The ONS has in the past carried out various improvements each decade and

it seems to be the kind of natural improvements that one would make there. It has broadly all the pros and cons of the existing census model. We did not dwell on that very long.

The alternative, which is the ONS's idea about how one could bring in administrative data into the census to replace some of the former ways of collecting data, is a much more radical option. It is one we looked at a lot. In some ways it is a slightly unambitious option, in that it is only looking to the administrative data for population counts—the numbers of people within age group, sex group, and areas. At this stage, it is not attempting to make use of other information, such as income from tax records, or the administrative data that some users think would make this option much more attractive. It is basically sticking with the administrative data for population counts, and then having this large, rolling, annual survey for all the other information. That rolling survey would be a sample survey. The proposal is for 4% per annum. A lot of the debate in the consultation has been about the limitations that would imply for this small, granular estimate—the small areas or small population estimates. There is a clear trade-off in the choice between that option and the traditional census option, in the sense that it would offer much timelier, annual data, but it would not offer the finest granularity that the census option does. That has been very much a topic in the consultation and I guess you will hear more about the pros and cons there.

Chair: That was a very helpful and long answer, but thank you very much for that.

Professor Skinner: I apologise. Yes, I could carry on, but—

Q49 Chair: I am sure you could. Are the other two witnesses happy with the quality of this consultation document? Does it deal with the right issues?

John Pullinger: Yes, I think it deals with the right issues. One thing I would note is that they say quite clearly in the consultation that it potentially understates the benefits of either options, in a sense, because it does not capture examples such as people like me as an individual in the community, or organisations and individuals that would not necessarily even know they are relying on census data when they are using numbers. The benefits described in here are likely to be understated.

The other thing that it does not do, which Professor Skinner mentioned, is that the administrative data option has been relatively unambitious. They have constrained themselves, in some way, to focus on the population by age and sex. To understand what would happen if they actually went down that route we need a much clearer appreciation of what individual users would actually do, faced with a different bundle of information. The timeliness would almost certainly be better, but how you would fill in the gaps—how someone like me, trying to make a case to a local council, would actually get hold of the data that was necessary, and what it would look like—I find very difficult to understand from this.

Q50 Chair: Under this system, would local authorities not, for example, actually become more responsible for collecting their own data about their own population?

John Pullinger: Potentially, yes, but that would then be another cost.

Chair: It would.

John Pullinger: You would also risk the comparability questions. In my point about what works, it would be difficult to compare Manchester and Liverpool if each local authority had done things in their own way. I know one of your other witnesses later has done some really interesting and innovative work in a number of particularly difficult London settings.

Q51 Chair: How much are local authorities lobbying in favour of maintaining the census?

John Pullinger: I have not seen their responses to the ONS's consultation, but my understanding is: strongly. I know you have one or more of them coming later. I have no reason to think they are not a representative group.

Q52 Chair: How satisfied are we that this really is a binding question about whether we carry on with the census or choose to use administrative data? Many of your answers seem to suggest that actually you need a combination of the two; it is not a binary question.

Professor Skinner: That is correct.

Professor Falkingham: I think that is right.

Q53 Chair: Then why are we not happy with the way this consultation is being conducted? Why have they given us a binary question in the consultation?

John Pullinger: Certainly, what they have said to me is that they have not considered it to be binary. They have given an evaluation of the online census option as shorthand and an administrative option. They have not said it is either/or. I think that is right.

Professor Skinner: Yes. Can I comment? In terms of looking to the future, if one were to take the administrative option it would be such a radical change. Because the research that has been undertaken so far has not yet demonstrated that it would definitely be fully reliable in terms of the estimates that would be produced, there is the question—even if one wanted to go down that route—of how one would proceed. Here, I would agree with John's word "reckless". To ditch the census entirely, and to proceed with this administrative option when it has not yet been fully demonstrated, would be reckless. One needs some means of validating it. The most obvious way of validating whether the population estimates produced by the new system are working would be, in fact, to run a census in 2021 and to make the comparison.

That also would have a massive benefit in relation to the comment earlier about time series. One of the most important things about the census is comparing things over time. If you ditched the current model entirely and replaced it, let's say, by an administrative model in 2021, and if you looked at the changes between those two occasions you will not know to what extent that is actually a difference because you have just introduced entirely new methodology, or whether real change is going on over that period of time.

Q54 Chair: Does shifting from form filling to online data collection not constitute quite a substantial change of methodology?

Professor Skinner: It does constitute a change, yes.

Q55 Chair: What effect would that have?

Professor Skinner: A few people have studied it a lot. It is a big thing at the moment. There are some effects. Some of them are actually improving reliability because with computer-assisted form filling sometimes, you can get asked questions. I was surprised when I went to the Dutch Central Bureau of Statistics, which is viewed as one of the top places. They now think that it should be the default that you always collect data over the internet, and you have to make a case that there is some subject that is not suitable for the internet to run a survey not using the internet. There is lots of research going on and yes, you are quite right that there are significant methodological issues there. However, they are not as significant as the proposed shift to this second option involving the administrative data.

John Pullinger: You can calibrate it because you have got the online and the paper and you can make some assessments of the effects. It is harder when you are making such a big methodological jump.

Professor Falkingham: We can look at the 2011 census for lessons as well. There was an option to fill the census in online in 2011, so we have lessons from the last census that we can apply going forward.

Q56 Kelvin Hopkins: Only 16% chose that option in 2011.

Professor Falkingham: That is true. That is because we all have paper put through our letter box as well. If we had not had the paper put through, I think it would have been much higher. Of course, you have to do something for the people who are not connected to the internet.

Q57 Alun Cairns: Can I pursue this a bit further? Are your calls not just another demand to analyse the data in a different way, about the approaches, and also to kick into the long grass for another 10 years, so that we will have this debate in maybe four or five years, for the debate thereafter? As statisticians, this is something that you love to see, love to compare and contrast and, effectively, kicking it forward, maybe with a slightly different emphasis on the data collection model next time, will take us to the next level thereafter and keep you all in business and interest.

John Pullinger: I would say quite the reverse. The key message from this is that the ONS needs to make much more rapid progress with looking at the potential of the administrative data option. The Treasury Committee was calling for this in 2008 and a certain amount of work has been done, but we have not yet got to the stage where we could see what the numbers would look like. That seems to me a very urgent next step for us to actually look at the numbers that would come out from this proto-system, compare it and start asking the hard questions that Mr Flynn has been asking. Can you do this at a fraction of the cost and still get the numbers you want? Until we see some numbers, it is very difficult to judge that.

Q58 Kelvin Hopkins: Just to make Professor Skinner's point, if you are going to change from one system to another, you have to have one year when both are operating because you are rebasing. If you want to move forward, you will have no comparison at all if you just stop the census and then, 10 years on, go to an administrative system as had been suggested. You have to have that comparison to make long-term comparisons.

Professor Falkingham: It does sound like we are asking for both of them, but if you look at the systems in Europe that have actually moved, they have moved over a 20 to 30 year period, so you do need to run them alongside each other for a period of time. That is exactly as Mr Hopkins—

Q59 Chair: How much is your message really that we need a process of evolution, not revolution?

Professor Falkingham: That would be a fair comment.

Q60 Chair: How possible is it to have a smaller census with fewer questions as we grow the capability of administrative data? How possible is it to do that when the tendency has been to grow the number of questions in the census?

Professor Skinner: That is a long-term thing. As I was saying, the ONS has not even targeted introducing additional questions in 2021. In our review, we looked at some of the questions, and in health, for example, it is really quite hard to use some of the administrative systems to get the kind of variables that you would like from the census there. However, I think there are clearly steps. If one looks at what has happened in Europe, Finland is the most extreme, where each decade it moved further and further, until now it is entirely administrative-based. A point perhaps needs to be made about issues of legislation if one is going to introduce administrative data systems. There is talk about data-sharing legislation that would introduce co-operation between the supplying departments and the ONS. What would be really desirable is having some commitment in that to ongoing improvement of these kinds of statistical outputs.

Q61 Chair: My understanding is that the Statistics Act already allows the ONS to make requests for undertaking those sharing or linking operations, but few, if any, of those requests have actually been made. Is that correct?

John Pullinger: My understanding is that it is a quite cumbersome procedure to use at the moment.

Q62 Chair: Is that the excuse?

John Pullinger: That is the argument that I am told. I have not sought to do this myself, so I do not know. It is more than just legislation here; it is the ability to get under the bonnet and understand what the administrative data are showing. The most recent example you have been looking at was the police recorded crime statistics. If you were starting to use these administrative data, which have been collected for a different purpose, there has to be an opportunity for the National Statistician to be able to review the quality of the dataset and, ideally, to set some kind of standards so they can be brought together and compared. The legislation is an enabler and it needs to be slicker than what is currently in the Act, but it also needs to be coupled with some procedures that enable the

National Statistician to have some kind of role in the data that we are going to use to create the kind of census that we need.

Q63 Chair: Do we conclude from that that we cannot stop the decennial census unless we have legislation to deal with this administrative data linking question?

John Pullinger: That is certainly my belief.

Professor Falkingham: Yes. We also need to know more about the quality of the administrative data. We need to investigate that more. There is also the issue that a lot of administrative data is about individuals, but individuals live in families, which live in households. The census actually allows us to look at households and families, and individuals. It is not clear yet that all the administrative data would allow us to reconstruct households. You would get different information from different administrative datasets.

Q64 Paul Flynn: What changes have the Finns reported in the reliability and costs of their census?

John Pullinger: The quality is better and the cost is lower, but it has taken a very long time to get there. It should certainly be our aspiration both to improve quality and reduce cost. The Finns do this very cheaply.

Paul Flynn: I am sure they do. It sounds like an intelligent system, which most Scandinavians do have. We seem to be still stuck in the past, with garbage in and garbage out—just an idle thought.

Q65 Kelvin Hopkins: On a recent parliamentary visit to Finland, the Finns told us that immigration does not really happen in Finland because the language is very difficult and it is very cold. That is what they told us. If they have a very small level of immigration, and very small population movements, doing a census in Finland is a good deal easier than in Britain.

John Pullinger: That is true.

Q66 Chair: Can I say I am getting rather a bleak picture? How fair would it be for us to report this as rather a bleak picture? The quality of the existing census is deteriorating, and the reliability of its statistics is less because it is more and more difficult to collect this data in a more and more complex society. More people are complaining about the quality of data, particularly at the local level, which it is actually meant to address. However, little preparation has been made to replicate the data by other means through administrative data. How well do you think the ONS is doing their job?

John Pullinger: I think that is a little harsh. If I might play it back a slightly different way, it is getting harder and harder to conduct a census in the United Kingdom in the traditional way. The ONS has been doing its utmost to deal with those challenges by estimating the people it struggles to count. At the same time, we have a huge growth in administrative data that is potentially available much more cheaply and quickly than a census would be able to provide. That points a way to the future, but we are not there yet. We would certainly wish to encourage the ONS to begin some experiments with the data it

can get hold of and to promote the cause for legislation to make it much easier for them to do it, so that by the time we get to 2021 we can have a lot more confidence. We will then rapidly be able to move away from the traditional method, which is harder, towards the administrative method, which is easier. That is less bleak.

Q67 Chair: Professor Skinner, for how long have all these issues been evident?

Professor Skinner: It has been a gradual process. I do not think I would take quite as bleak a line. It seems that the ONS has been aware of the undercount issues on the census and the things that it has done have improved them. That is particularly in terms of actually avoiding it in the first place, with the success they had in 2011, but also, given that some of it is going to exist, the methods for carrying out corrections. The level of concern in 2011 seems to have been rather less than in 2001. They have been doing a lot of innovative research on possibilities for administrative data. It is still at a relatively early stage. In an ideal world, some of that research may have been further on, but they have had the opportunity to capture administrative records in 2011 at the same time as the census, and that has been a really important data resource for carrying out some of the research. To some extent, the key starting point is when that data became available in 2011.

Q68 Chair: Professor Falkingham, how confident are you that this will be tackled with sufficient urgency?

Professor Falkingham: I am fairly confident. There is a significant research programme ongoing at the moment within the ONS around different methodological issues. They know what the problems are and they are working hard to find the solutions.

Q69 Chair: Finally, can I just ask all of you how well do you think Ministers understand this?

Professor Falkingham: It is a very complex issue. We have struggled today sometimes to try and explain some of the statistical intricacies. There will be some Ministers who understand it very well and others who understand it less well.

Q70 Chair: Do you think the Government has the capacity at Ministerial level to grip this and understand it? How confident are you that there is sufficient Ministerial knowledge and attention being given to these matters?

John Pullinger: That is a difficult question for you to put to us.

Chair: It is not, because you must have a view.

John Pullinger: The next stage is for the ONS to publish its results to its consultation, to which we have contributed, Professor Skinner particularly by doing a methodological evaluation of it. The test will be at that point when that comes out. As far as I see it, quite a clear and consistent view has come back from the user community, both inside and outside Government, about the way forward and the risks of particular alternatives. It is now clearly for Ministers to think about that and decide.

Q71 Chair: Yet the Minister of State for the Cabinet Office has already made it clear that he is against having another census.

John Pullinger: He starts from the position that the plethora of open data and administrative data ought to make it possible to do a census in a much cheaper and quicker way. We all agree with that. The test is one of practicality now.

Q72 Chair: Are there any other comments? Thank you very much. You have been very informative and I have certainly learnt a lot from this session. I am extremely grateful to you, thank you.

Examination of Witnesses

Witnesses: **Keith Dugmore**, Director of Demographics User Group, Demographic Decisions Ltd, **Sarah Henry**, Head of Intelligence and Performance, Manchester City Council, **Juliet Whitworth**, Research and Information Manager, LGA, **Piers Elias**, Demography and Modelling Officer, Tees Valley Unlimited, and **Professor Les Mayhew**, Professor of Statistics, City University, gave evidence.

Q73 Chair: Could I ask each of you in our second panel to identify yourselves for the record, please?

Sarah Henry: I am Sarah Henry. I am Head of Intelligence and Performance in Manchester City Council.

Juliet Whitworth: I am Juliet Whitworth. I am the Research and Information Manager at the Local Government Association.

Piers Elias: My name is Piers Elias. I work for Tees Valley Unlimited. I am a local government representative on the Census Advisory Group, and I also co-chair the Population Liaison Group that meets the ONS and central Government to discuss methodology and ensure local authority needs are catered for.

Keith Dugmore: My name is Keith Dugmore. I organise the Demographics User Group, which represents some large, commercial end-users of census data.

Q74 Chair: Thank you. I should add that we are also waiting for Professor Les Mayhew, the Professor of Statistics at the Cass Business School. Unfortunately, he is launching his book today. I will give that a plug for him. This is unfortunate for us, but he will be turning up and I will fire some questions at him when he arrives. Before we go on, I have noticed you all listening to the previous session. Are you burning to say anything in reaction to what you have already heard?

Keith Dugmore: I have one quick point on what Mr Flynn raised about the merits of alternative methods. I am completely with him on the possibility of using administrative data to a greater extent, but when it comes down to samples and sample survey, those of us who are concerned about statistics in very small areas see this as a great danger. It could render a census useless if we could not get statistics for very small areas.

Piers Elias: I have two points. First, the 2011 census was better than the 2001 census. We are getting better data for two reasons. One is that the address register was properly checked, updated and cleaned by local government and by the ONS. Secondly, the response rates from the hardest-to-count authorities were much higher than in 2001. For those two reasons, it is wrong to say that 2011 was worse than previous censuses.

Chair: I can see assent from other members of the panel. You do not need to say anything if you do not need to.

Sarah Henry: Yes, I agree.

Q75 Kelvin Hopkins: You have heard the previous session and I asked questions about the value of the census. It seems to me from earlier comments that for every pound spent on the census, the private sector alone benefits by six times that amount every year. Does that not suggest there is a value?

Sarah Henry: I agree that there is huge value in the census. It is very difficult to put an exact pound on that value, but looking back at decisions that are made at a local authority level, I think that the census sits implicitly or explicitly behind every decision that we make, with a focus on operational decisions as well as big strategic decisions. All our ward councillors are very aware of the place that they represent, the people whom they represent and the intelligence that they get to inform them. That comes from the census—not only the census, but it comes from the census.

Q76 Kelvin Hopkins: We have heard that the census has improved over the last decade. That is one good thing.

Sarah Henry: Certainly from our perspective.

Q77 Kelvin Hopkins: Is there not an argument that the alternative, administrative methods of collecting data should be an addition to the census rather than instead of the census?

Sarah Henry: Certainly now, they should be in addition to. I agree with the previous panel that we are nowhere near replacing the census with administrative data yet.

Piers Elias: I would just add that the ONS use a lot of administrative data in their population estimates. They do actually link individual data between the GP patient register and the higher education data on students to try to improve their methodology on migration from year-to-year, so it is already happening. We would encourage the ONS to continue to do that. What we do not get is the characteristics that are sitting on those admin files and we would be very interested to hear of the ONS doing more work on digging those out.

Q78 Kelvin Hopkins: My first two questions were leading questions, but now, an open-ended question to each of you: what do you personally think is the value of the census?

Keith Dugmore: You mentioned commercial views and I will just put our perspective on that. I have worked in local authorities before, but in the commercial world a lot of

companies use census data to try to understand local markets. They use this for locating stores, in particular. Just look at the extent to which companies like Sainsbury's have been opening stores over recent years, or Whitbread opening Costa Coffee. All these things involve a lot of investment and having reliable small area statistics is vital for those sorts of decisions. There is also the question of what they put in the stores and how they vary the offering according to local markets. A third area, which has been mentioned already, is organising market research sample surveys. Again, the major companies there rely heavily on census data to target particular groups.

When one thinks of the opening of supermarkets, for example, a large one might be an investment of £25 million to £30 million; smaller ones might be £5 million. Even a few companies being involved in that soon add up to big numbers. When you think that we know that many thousands of companies use census data, the numbers become potentially very big indeed.

Juliet Whitworth: I would add that local authorities are major users of particularly the small-area data. This is the case below local authority level at ward level, and even below that. High-level information is useful, but you can miss the granularity. You can think that you have got an area that is full of owner-occupiers, for example, and miss pockets of private-rented housing or social housing. This information and this granularity are really essential for local authorities to be able to target their resources, and, as Sarah said, to be able to make not only strategic decisions but also day-to-day decisions. I would argue that it allows delivery of central government policy as well, such as with things like the localism agenda, doing neighbourhood development plans. They often do not fit within normal administrative boundaries, and because of the level of detail that is available from the census, you can build up from really small units to understand those neighbourhoods where work is going on. It is really crucial.

Local government has perhaps struggled to place a numerical value on that small-area data. That is not because it does not have a value, but more because it is quite difficult to measure that targeting of resources and that sort of thing.

Sarah Henry: If we did not have the census and we had to carry out endless surveys, that would not only be very time-consuming, it would be very resource intensive for us. We need to take that into consideration. It is helpful to distinguish between the count that we get from the census and the rest of the population characteristics that we get from it. Both are extremely important, especially for democratic accountability at very local levels. Ward members need to understand their population and the people they represent; they need to understand households and place. That is really important for planning for growth and reform.

Piers Elias: The other thing the census gives us that would be very hard to replace is flow data. For example, travel-to-work data is very important for local authorities. It helps with transport planning; it helps businesses decide whether the catchment area of the place they are going to invest in has a big enough pool of resources or skills. We would miss that without the census.

Q79 Kelvin Hopkins: On the point of granularity, I represent Luton, which has a large Asian population. In the past, people have talked about Asians in general terms. In

fact, in Luton, we have overwhelmingly Kashmiri and Bangladeshi, whereas in Leicester they are Gujaratis from East Africa, professional-background Indians. These are very different populations generalised as Asian and the granularity is really important.

Sarah Henry: Manchester has the same issue.

Q80 Paul Flynn: How many times do you think the manager of my local Sainsbury's would consult census figures for his decisions?

Keith Dugmore: The manager of a local Sainsbury's?

Paul Flynn: Yes, my local Sainsbury's.

Keith Dugmore: I do not think that he would refer to them directly, but he would be talking to the site location team at Sainsbury's headquarters, who would inform him what the local market was and how it differed from others.

Q81 Paul Flynn: The local Sainsbury's came to my constituency 35 years ago and they relocated to a bigger site. The decisions they take are based on what the customers are buying. What they will order for next week depends on what they have sold in the previous week. Are you serious about saying that the census is something that they look to guide them in their daily decisions?

Keith Dugmore: Yes. There are two elements to that. One is changing the estate of stores. Companies such as Sainsbury's, which is a good example, have grown their number of stores from roughly 500 to about 1,000 in the last decade. They have needed to decide where to put the new ones. They do not have data.

Q82 Paul Flynn: How could the census help? I am genuinely puzzled by what you are saying. They will do it on the basis of population they know, on the profitability of the shops, and on the customer and the customer's taste. That is not in the census.

Keith Dugmore: They are concerned in getting information about the numbers of people and types of people in local areas and also questions of access; they are interested in transport and so on. They will then look at the market potential in different parts of urban areas and decide that Balham is better than Tooting or whatever. When they are locating new stores, those are the sorts of decisions that are made. I mentioned Costa Coffee, and they are very interested in where people are working. It is not so much where they live but where they are going to be at lunch time. Boots is another example of people popping out to buy things at lunch time. The information about where people work that we get from the census is absolutely vital for locating stores.

Q83 Paul Flynn: The big supermarkets do their own market research. They have their own experience of what is selling in their shops, what people are coming in and it is based on that. It is not a decision taken by a census; that was 10 years ago.

Keith Dugmore: I can say definitely that that is not the case. They certainly have their own market research—

Q84 Paul Flynn: If Sainsbury's in Manchester are making a decision, should they base it on what the census says or what the local authority says of the population of Manchester? There is a lot of difference between the two.

Sarah Henry: I do not think it is either/or. When you make those kinds of business decisions, you take everything into account, including the census and including your own market intelligence. You also use various products that are on the market that are based on the census, such as the predictive modelling that Acorn and MOSAIC provide. It is very difficult to disentangle the census from all the other bits of intelligence that we use for various decisions, and that local businesses use for their decisions. The census is the bedrock. Even if they do not think they are referring to it immediately, they might still be referring to it.

Q85 Paul Flynn: We know there are major imperfections in the census: 5% of the population do not take part in it. Many of these are migrants who might have language problems or who do not trust authority; many of these might be elderly people who feel they cannot cope with it. There are all kinds of other reasons. We already know there are those people trying to manipulate it. We already have a lot of garbage going into census forms. Is it not inevitable that garbage will come out of the other end?

Sarah Henry: No, because there are statistical methodologies that we use to improve them. That kind of bias applies to all the other methodologies that are being proposed. Surveys are worse than the census in that respect, because there is self-selection in whether you decide to participate in the survey or not.

Q86 Paul Flynn: This is another, rather elaborate way of describing guess-work that comes up as an element of that figure. Have any of you examined what is happened in Finland? They have gone on to an administrative system there that we hear is cheaper, timelier, more efficient and more scientific, by the sound of it. Why are we lagging behind in the past?

Sarah Henry: As we heard from the previous panel, it is easier to carry out a census there.

Q87 Paul Flynn: It is always easier to be inert and carry on with the system that you have been using for a long time and that people are reluctant to change. Is that what it is? Does Finland have a system that is more scientific, cheaper and more reliable? Yes or no?

Juliet Whitworth: I do not know a lot about Finland's system in detail, but the other issue is that they did not arrive there overnight. It took some time. We are not saying that we should not move to administrative records. That is very likely to be what we aspire to do. What we are saying is that we want to make sure that when we make that move, the administrative method is as good as it can be and gives us what we need, and that we do not just junk something before we have got an administrative system that really works.

Q88 Paul Flynn: Could I ask Professor Mayhew, who has caught his breath now—?

Chair: Can I just ask Mr Elias to make a comment?

Piers Elias: You were saying that the ONS guessed the last 5% of the people who do not fill in their forms. First, you have got the 94% who have filled in their forms. The other

5% are built up by comparing with administrative data sources. Some of them are actually genuine people whom they can find and they put them back into the census. It is not all guesswork. A lot of it is done by maths and matching with administrative data.

Paul Flynn: It is elevated guesswork; I will give you that.

Chair: That is what administrative data is.

Paul Flynn: Yes, indeed. It is based on a large, reliable sample. Can I ask Professor Mayhew and say that I raised these points in 2005 as a member of the Committee about the upcoming census? I thought it was a biblical solution rather than one that matched modern science.

Q89 Chair: Welcome, Professor Mayhew. Could just identify yourself for the record before you answer.

Professor Mayhew: I am Professor Les Mayhew from Cass Business School, City University, London, Faculty of Actuarial Science and Insurance.

Q90 Paul Flynn: Is there a better way of collecting the information we require from the census, rather than having this 10-year attempt to interview everybody?

Professor Mayhew: I am an advocate of an administrative-based system. I have a small company that has been estimating population at the local authority level, up to the size of cities as large as Birmingham, for about 30 years.

Q91 Paul Flynn: How does the accuracy of your figures compare with the census?

Professor Mayhew: Very well. It depends where you are. If you are in London, we certainly found 4% to 5% more people when we compared our figures with the then ONS mid-year estimates. We were commissioned by the Olympic boroughs to do this because they were concerned that the census would not adequately count their populations.

Q92 Paul Flynn: What was the comparable cost?

Professor Mayhew: The comparable cost for six boroughs was about £200,000.

Q93 Paul Flynn: We know the last census was £482 million and there is a lot of information on that. Would you agree that the last census will probably be the last census? Would it not be an improvement to have a continuing census going on with a tiny sample, compared to the 100% sample that we seek now? It would give us information that would be timely, of good value and more reliable.

Professor Mayhew: You have to look at this in layers. There are the data available in administrative systems. There is absolutely no doubt that central Government data systems could provide a good population count and even a household count if you did it properly. Whether you would get all the attributes you needed that you get in the census, I doubt. However, you would also get other attributes that you do not currently get, such as whether a person is on income support or not. I also think that there are data that you will continue to need to collect via surveys. These are softer figures, like how many hours a

week you spend caring for an elderly person and that sort of thing. Surveys will remain the appropriate way for doing that.

The other area that I do not think has been mentioned is the enormous amount of data assets that are held at a local authority level that they use for a multitude of things: locating health centres, assessing child care, or looking at troubled families. There is a whole multitude of applications there for which the census does not really give you any adequate information at the level of granularity you need. I think there is a model that says, "Yes, the centre does the big population count, but it then tags people to local areas". Then local areas can add or append their own attributes to that data for the purposes they need at that point in time. At the moment they cannot do that, because the granularity is not sufficient; the data are out of date; geographical units are wrong; and so on and so forth. You have got the beginnings of a model there which involves local and national collaboration, overlaid with a survey to get the softer information that is more difficult to collect from administrative systems.

Q94 Paul Flynn: Before you arrived, Mr Dugmore was saying that one of the advantages of the 10-year census is that it did have granularity, and there were small areas in which it showed the results that perhaps would not be apparent. Would it not be better to have targeted investigations into those particular areas to provide a more accurate picture, and not one that came up by accident in the statistics?

Professor Mayhew: I certainly think that the proposed administrative approach needs to be road-tested in some big areas. Different models need to be tested that involve combinations of local and central input. However, I do think that if the centre does not provide local areas with the statistics they need for their local usage, they will create their own statistical systems that suit their needs. In fact, I know many local authorities are already doing that.

Sarah Henry: Yes. I agree with that, but if we lose the census, we will lose some of the benchmarking. There is a long way to go in order to agree how we carry out any administrative option. Personally, I do not think there is enough time between now and 2021, for sure. The legislation needs to be thought out carefully; it needs to be enforceable; different Government Departments would have to have duties around population statistics set out in law, which they do not have at the moment. The other point is that in order to do the kind of work that Professor Mayhew is talking about, data need to be matched. In order for it to be matched, it needs to be identifiable. The experiments that have been done to date by the ONS did not include that exercise.

Q95 Paul Flynn: These arguments were the ones presented in 2005, very much from people who are comfortable with the present system and the status quo, and do not want to change or are afraid to change. We have a system that is far less than perfect. Nobody believes that the fourth biggest religion in Britain is based on Star Wars, and there are other wildly inaccurate figures for Westminster and Manchester. We cannot talk about this as though it is the gold standard of censuses. We seem to hear that better systems are coming out at a fraction of the value from Finland and other places. Is that not true?

Sarah Henry: Nobody is arguing that the census is perfect, and in 2005 Manchester City Council would for sure have had a lot more to argue against it. In 2011, though, to the

ONS's credit, they learnt a lot of lessons from 2001 and they worked with us in partnership in a way that we think can be built on to improve the next censuses even further.

Q96 Paul Flynn: Could you give me an example of something that came out of the census we would not know in any other way?

Sarah Henry: Yes. For example, they used us in order to generate the address base, which they did not do in 2001. In 2011 all local authorities contributed to that, and the address base used for the census was far improved. Come 2021, we believe that we could help in ways that we did not in the past. For example, we could go round to people's houses all the time. We can help collect forms or, hopefully in 2021, we will be going round with mobile devices and looking at how people fill in their forms online. We can make the census better and cheaper that way, if we build on the partnership.

Q97 Paul Flynn: Why do we need a census to tell us how many people are on income support?

Sarah Henry: I do not think we ask that question—

Paul Flynn: No, it was something Professor Mayhew mentioned.

Sarah Henry: —But about unemployment and the kind of skills people have. That kind of information is not captured at the moment by DWP.

Piers Elias: I will just to come back to Professor Mayhew's point about local authorities using administrative data sources that they hold. We do use a lot of council tax data, Land and Property Gazetteer data, and, for example, the electoral register, which has recorded nationality since 2006. We do most of that as a proxy for international migration to the area. There are the two issues with it though. One is that it does not cover the whole population; it only covers the 17-plus population. Secondly, according to Audit Commission figures, 10% of us are not on electoral registers. These admin data sources are potentially very useful. If we can tighten up the legislation that means people have to re-sign up when they move or something like that, that would be really useful.

On the point about the address register, there was no legislation after the 2011 census to allow it to be continued to be used because of intellectual property rights arguments between Ordnance Survey, Intelligent Addressing, and the Local Government Association. If you could get rid of that and put some legislation that allows us to keep that cleaned address base to start with, we have an address register.

Professor Mayhew: I just wanted to say that whilst I agree with Piers, the data that are available locally include the school pupil census; it includes other data about older people, social care, and so on and so forth. Our methodology combines all these datasets together alongside the GP register. We have a system for sifting through and confirming every individual in the population by the address that they live in so that we can confirm that person on the basis of several pieces of evidence. Secondly, we can assign them with a household, which we can also find—is it a lone parent household? Is it an older person living alone? Is it a three-generational household? All those things become possible if you combine these datasets in the right way and it all can be made fairly automatic.

The challenge is to actually clear the bureaucratic undergrowth to enable these things to happen automatically, instead of having all these disputes between different data owners, sometimes under the guise of data protection and things like that. All these things could be done in a data safe haven. People can be accredited to use this kind of data, and the outputs that come from these processes could be fully de-identified and anonymised as well. That is something that we do very regularly, so we know it is possible.

Keith Dugmore: Can I pick up two points? For the local authorities themselves, there is probably a lot of interest and a lot of strength in using their local data for the census. Commercial companies, though, are particularly interested in getting a consistent national view, so I think it is very important that some of that data is produced from the centre, rather than having local variations. Dare I say it, national view actually extends to the United Kingdom as a whole. There is an interest in not only getting England and Wales, but adding Scotland and Northern Ireland so that one can look across markets across the entire UK.

On the second point about the National Address Gazetteer and the fact that it was created for 2011 and has now become a commercial product through Ordnance Survey, this is one example of where it is important that Government puts its foot on the accelerator to try to improve access to the address file, and some of the Government Departments' administrative files. Otherwise, there will be a handwringing in the next five years with people saying, "Well, we would like to do it by 2021, but it is all very difficult", and another decade passes. The ONS has made a lot of progress in their research for Beyond 2011, looking at the Department for Work and Pensions, the National Health Service and so on, but it needs to be propelled more quickly in the next couple of years. Then we might see where we are as we near the date of the next census.

Sarah Henry: I agree that the matching exercises are extremely useful. In Manchester, in preparation for 2011, we actually did a similar exercise using our own administrative data. Admittedly we did not have access to everything that the ONS would have access to, but we did that. We ended up with a range that the actual census figure fell in, but at the higher end, so there is something that needs to be considered around people who are not listed on any of the registers. For example, we found out that there were 20,000 homes in the city centre of Manchester where nobody was registered with a GP. The service was not nearly as universal as everybody thought it was. When you come to urban areas where there is a lot of change in the population, it is quite difficult to enumerate. You need a foundation that is consistent nationally, but we think that for big urban areas—Manchester and Westminster are two good examples—we need to think very carefully about how we use administrative datasets.

Q98 Greg Mulholland: If I could just ask this to you specifically, Sarah Henry. You have clarified to some extent the concerns that Manchester had. It is very much one of the councils where there were problems—45,000 people were missed in the 2001 census and Manchester City Council raised concerns. Westminster City Council still has concerns. We heard about those concerns last year. Are you now saying that Manchester City Council is entirely happy with the reliability and the accuracy of the census, even though Westminster still has exactly the same concerns?

Sarah Henry: We were much happier in 2011 than we were in 2001, and we worked very hard with the ONS to achieve that. I am aware that Westminster was less happy. Obviously, their population seems to be harder to enumerate, but I still think that if we build on the partnership, we can make the census better. That is my point. I am definitely not saying that we have got to the point where we have made it perfect for everybody, but it is not about throwing the baby out with the bathwater.

Professor Mayhew: It is very easy to select Manchester and Westminster as outliers. They have been known for a long time. Let me give you an example in the other direction. I have worked on and off in Newham for about seven or eight years. When we last went there, the mid-year estimates said there were 243,000 people living in Newham. When we did our administrative work in 2011, in time to coincide with the census, we identified over 300,000 people. When the census finally published its results, they also found 308,000 people in Newham, so for several years there has been this discrepancy of over 60,000 people in Newham that has affected schools planning, service planning and a whole host of other things. They were particularly keen to get my view, to produce a different perspective, a different evidential approach, and come to that figure. In fact, we found that in several locations in London, none as big as Newham. Newham was a very special case, with lots of transient in-migration and other things, but we found quite substantial discrepancies in a host of other London boroughs.

Q99 Chair: So it would be inaccurate to say that it was the census alone that identified the miscounting of the population in Newham.

Professor Mayhew: They started from a baseline in 2001 that may have been a false baseline. Other people can comment on that. Then they used their methodology, adding births, subtracting deaths and adding net migration, year-on-year, up to 2011. The figure they got to was 243,000, but the actual count of population in Newham was over 300,000.

Q100 Chair: That suggests that depending upon administrative data is what led them to the wrong estimate of their population.

Professor Mayhew: Potentially.

Q101 Paul Flynn: What was the difference with yours? How did you achieve your figure of 300,000?

Professor Mayhew: We combined all the datasets that I have mentioned. We had our rules for deciding whether that person was genuine and lived at that address. We came to that figure and, interestingly, the 2011 census came to an almost identical figure, but they had been working on a completely different assumption in Newham for several years previously because of the false baseline in 2001.

Q102 Chair: Even if you got your estimate correct, it suggests that the census was the only qualitative assessment that could be produced that would verify your figure or not.

Professor Mayhew: That is true.

Q103 Chair: How can we then consider getting rid of the census?

Professor Mayhew: There is one thing we do that I have not mentioned. We have used random checks on households to see who is actually living there and compared that with our own figures on a sample basis. I agree though. I am not saying that the 2011 census would be an independent check or not. In fact, we were the independent check on the census. By the way, we produced our figures within six months, compared with something like 18 months.

Q104 Chair: My understanding is that your position is that the census should be scrapped?

Professor Mayhew: My position is that there are much better ways of doing it, and we would save a hell of a lot of money if we did.

Q105 Chair: On the other hand, you have verified that, if we had not had the census, we would not know the true population of Newham because we would not have been able to cross-check your estimates. There is no other benchmark your estimates could be cross-checked with.

Professor Mayhew: Well, suppose you did my exercise every year or every quarter, which would be feasible within the envelope of money that you currently spend on the census.

Q106 Chair: Sorry, can you just answer the question, please?

Professor Mayhew: You would have ongoing checks and verifications on who was inhabiting particular households in your patch, and you could analyse that data and use it as partial confirmation of what was going on in your area. I am convinced that this is a cheaper and more accurate alternative to the census.

Q107 Paul Flynn: What percentage of the households did you contact, and how many do you think you would need to contact to have a continuous picture?

Professor Mayhew: In the one place that we did this, which I think was the London Borough of Brent, we contacted something like 1,000 people who were on what was then called a “citizen’s jury list”. We are a small organisation.

Q108 Paul Flynn: That would be a tiny percentage of the total population.

Professor Mayhew: Gallup and other organisations use small samples. You choose your confidence limits and select your sample accordingly.

Q109 Chair: We have the delight, on this panel, of having people with different views and we would like to hear them.

Piers Elias: On the GP patient register nationally, at the 2011 census, it was found to be about 4% overinflated. It ranged from plus 15% to minus 15% overall for local authorities. If you go into the age groups and look at the 20 to 30-year-olds, you find huge disparities of up to 50% differences in counts for people on GP registers compared with the census. That is because students are not very good at re-registering when they leave

university. It is because of the migration. People might come to the country, not register where they first live, and then move.

When an international migrant comes into the country, they get something called a “Flag 4”. Newham actually raised that idea of using Flag 4 status as a way of counting international migration. It is used in the current method of allocating international migrants at a local authority level. We have to have a benchmark to check the admin sources. Without that, we can end up going in the wrong direction, and that has issues of fairness for funding for local authorities; some are getting too much, and some are getting too little. The ones who get too little tend to shout the loudest.

Sarah Henry: In the interest of balance, the census is not perfect; we have determined that, but the administrative datasets are not perfect either. That is why we need to carry on working on both methodologies. It is still worth having the validation every 10 years, in order to make sure that error is not creeping in if there is a significant shift. It also gives us added value of being able to cross-reference and cross-tabulate lots of data that are collected at one single point. You cannot do that in a survey, otherwise the surveys will be too long and nobody will participate. The statistical confidence that you cite assumes that you have randomly sampled and that everybody has responded. That does not happen. There is self-selection in surveys. Neither solution is perfect: keep that balance.

Q110 Paul Flynn: How do you explain that a survey that attempted to interview 100% of the population came up with a figure that was wildly inaccurate, and a survey that interviewed a tiny fraction of 1%, conducted by Mr Mayhew, produced a figure that was accurate?

Sarah Henry: I have not looked at the detail, but I am assuming you included lots of other variables in your estimate, not just one. You also include far fewer questions in that survey than are in the census.

Professor Mayhew: It depends whether you are just trying to verify the population count, or whether you are trying to collect 200 other variables from each individual. There is a substantial practical difference between those positions.

Q111 Paul Flynn: For the enlightenment of the nation, Professor Mayhew, would you tell us what the title of your book is, which you have just published today, I understand? I think we need to read it.

Professor Mayhew: That was on paying for long-term care.

Paul Flynn: Oh, was it?

Professor Mayhew: It was carried out with Demos. I can give you a copy now, if you want. I have just rushed from the launch of that to here. I am talking about my other passion, which is population statistics and better use of them.

Q112 Chair: We are very grateful you got here. Juliet Whitworth, do you have a comment?

Juliet Whitworth: It has been picked up since, really. Sarah made the point that, although the census is not perfect, and the 10-year census is not perfect, the administrative method is not proven to be perfect yet either. We do not disagree in the sense that we are saying

“never” to administrative data. That seems a really sensible way to go. Local government would really like more frequent data, if that was possible, and administrative data is a good way of getting that. What we are saying is that we should not rush into it before we know it is good enough. Running the two in parallel in 2021 would allow us to assess that.

Keith Dugmore: I agree wholeheartedly with what Juliet has just said. One thing about administrative data we may be overlooking is that it can offer us new topics that we do not get at the moment. It is not just a matter of replacing questions in the census. There has been a long interest in information about income that has never been asked in the census. When we look at HMRC as a data source, we think, “Wouldn’t it be good to be able to harvest some data there, and produce aggregates and anonymous statistics for small areas?” There are some distinct pluses with admin data, as well as just substitutions.

Q113 Chair: Again, I am drawn to a question I asked the other panel. Is it really helpful of the ONS to distil this into a binary question between two options, either administrative data or a census, when in fact you all seem to be saying that you need both?

Keith Dugmore: I agree with the point you are making. When I read the consultation document, my feeling was that people are bound to say, “Do I vote for A or B?”, and it risks not thinking about a third option. In particular, the second option, which was described as “administrative data plus surveys”, really is primarily surveys, plus a little bit of administrative data. It would not give us anything much at a small-area level. Some of us in the commercial world would urge the ONS to continue the good work and do it more quickly, and to try to get more access to administrative files swiftly in the next year or two, which probably involves legislation on data sharing.

Q114 Chair: Is the binary question helpful, Mr Mayhew?

Professor Mayhew: What we might be looking at is a hybrid, where the main system is administrative-based, and you produce frequent—it is open to question what the frequency should be—population counts down to the local level. The local level can append and attach its own data. Some of that is centralised. Centrally produced data will have attributes on it such as those that have been suggested. Then you might carry out surveys, a bit like the one that I have done, to verify that accuracy, but you would do it on a smaller scale, and you would work in more difficult areas where counting population was slightly more problematic.

Q115 Chair: This would involve not a decennial census, but perhaps a much more regular but much more limited census.

Professor Mayhew: Yes. You would just rotate it and do it on a directed basis to areas that you were concerned with.

Q116 Chair: I would like to have the other members of the panel on this question. How helpful is it to have a binary question in this consultation?

Sarah Henry: It is not helpful, and that is why we did not answer the question in that way, but responded that we would prefer a hybrid.

Piers Elias: The ideas for a third way have only just started coming out, really, towards the end of the consultation. There were some ideas from the British Academy that Professor Sir Ian Diamond has suggested, where you have an online census and then you use those records for individuals to update in the future. That is either as their own updatable piece of information they update when a certain event happens in their life, like they move home, or every three years or whatever.

Juliet Whitworth: We also did not respond in a binary way, and responded to say “both” in the first instance.

Q117 Chair: Very interesting.

Q118 Paul Flynn: Can you give us some idea of how much business relies on the census, now out of date? It will be 2023 before we get information on the next one. Given the day-to-day reports that are put out on a weekly, monthly, annual basis by the ONS anyway, do they not rely most on the regular reports that come out and are much more valuable?

Keith Dugmore: The quick answer is they do not. The particular value of the census is the fact that information is for small areas, and that is what they need to make their decisions about local store catchments and so on. The ONS’s publications on national trends or regional trends may be useful in other parts of their businesses, but when it comes down to investment in local stores, that is of no use, really. It is the fine geography of the census that is valuable to them. If one went back two or three years, before the 2011 census came out, the 2001 census was still being used for store location decisions, albeit 10 years old, because there was nothing better.

Paul Flynn: There was in Newham.

Piers Elias: A lot of characteristics do not change very quickly. For example, the pool of skills you have in your area is going to be very similar from one year to the next. If you have a survey that is trying to gather data over five years, for example, which is what we would need to get small-area data, you have 10%—or 14% in London—of people moving every year, so it is going to be a very confusing statistic. At least a census is a snapshot; it is a point in time, and we can understand that much more easily.

Q119 Chair: Supposing we went for option two, as it is described in the consultation paper, what would we lose that cannot be replicated in the way they described?

Piers Elias: We lose flow data, so that is travel to work and migration statistics. We lose a lot of information for small parishes. In Suffolk, it is almost a complete wipe-out in terms of characteristic information. You get population counts, but nothing else. Just about every single parish in Suffolk is below the threshold needed. That is the same across several rural authorities

Q120 Chair: Do all members of the panel agree with that?

Professor Mayhew: Not completely, because I think you gain and what you gain could be argued to be more than what you would lose.

Q121 Chair: You think we can do without all that data that Mr Elias described.

Professor Mayhew: I am considering what is relevant for local authorities in their day-to-day work. What gets them excited is how many people will turn up at school on the first day of term, how many people will need social care in the future, how many post offices there are in the area and whether these are accessible to vulnerable groups.

Q122 Chair: Your argument is, essentially, that the data you lose, which Mr Elias has described, is actually not data that you need.

Professor Mayhew: There are uses for it. I am just saying that the local authority users have quite different perspectives and needs. That is why I argue that they should be enabled to append their own attribute data to national administrative data produced at the centre. That would overcome that problem. There is not one spanner that fits all nuts, basically.

Q123 Chair: You envisage replacing the national census, to a degree, with a whole lot of local authority censuses.

Professor Mayhew: No. The vision is that the centre would produce the core data, counting the population and certain key attributes that are available from administrative systems. Local authorities would be able to append their own data to that data, as and when they need it. Then the other part of it, the softer data—the issues about religion and so on and so forth—could be picked up in surveys. You would lose something at a local-area level, and I agree with Piers on that, but you have to consider, in the great scheme of things, what the priority is and which is the most important data needed by users. Local authorities have been a forgotten element in this whole debate for many years. They do not voice their views as loudly as they might. Often, we hear the academic perspective or other perspectives instead, which are valid, but do not fulfil this need.

Q124 Chair: Mr Dugmore, it could be argued that the users you represent, if they really want this data, should collect it for themselves, and pay the cost of collecting it for themselves.

Keith Dugmore: I have two things on that. Firstly, I must pick up the point that the ONS's second option in the consultation involves some admin data, but would only provide statistics by age and sex at small-area level, and it does not provide anything else. It is important that is recognised when people are looking at the consultation. It is a very limited admin option.

Chair: I see Professor Mayhew nodding at that.

Professor Mayhew: Yes, I think they ought to consider adding a lot more attributes to that data. I completely agree with that. What they are proposing is very limited. They cannot even construct households from the way they are doing it because the data is anonymised.

Q125 Chair: Is option two an Aunt Sally option?

Professor Mayhew: Possibly—it is an unambitious option.

Keith Dugmore: It is a very cautious option. It really is.

Q126 Chair: You were going add something else, Mr Dugmore.

Keith Dugmore: On the point as to whether anybody else could do it better or companies do it themselves, the great attraction of having something centrally organised through Government is that you get universal, even coverage across the country, and it has many benefits for non-commercial users as well—the great British public. The Government would either do that as it always has done as a traditional census or mobilise its increasingly large administrative data resources to come up, again, with something that provides good coverage across all areas of Britain. Nobody else can do it.

Sarah Henry: Arguably, they have paid for the census to be done in the most efficient way, through the taxes that they pay. The Government is spending their taxes. It would be extremely inefficient, across the whole of the industry, for people to produce their own censuses.

Chair: That is an extremely powerful point. I am almost running out of questions here, because we have dealt with this so comprehensively.

Q127 Paul Flynn: Has any other member of the panel apart from Professor Mayhew any comment to make about why it is thought to be essential to spend the ratepayers' money on Professor Mayhew's investigations? Were they wasting their money, or is there a really serious defect in the census material?

Piers Elias: The reason is usually to do with local authority funding. They tend to be local authorities who are undercounted.

Q128 Paul Flynn: That is very tight, and I am sure they do not have money to waste on this, but they took the serious decision because the census material on which they could take major decisions was not reliable. In Newham, they would have been taking the decisions that were hugely inaccurate, based on an inaccurate thing. It does expose the census as being far from perfect. There is this idea that any census can be perfect, which it is not going to be, but here we have a clear example of a major defect in the census.

Piers Elias: It is not a major defect in the census.

Paul Flynn: 243,000 population compared with 300,000.

Piers Elias: The ONS use the GP patient register that Professor Mayhew uses to count the population in the mid-year estimates, so it is obviously a problem with the stock of the GP patient register at the start of 2001. That is not an issue to do with the population being wrong in the census.

Q129 Paul Flynn: The result was so granularly accurate that it missed out 50,000 people in Newham alone. How can you defend that?

Piers Elias: It could have been that the base from 2001 was too low.

Sarah Henry: The other thing that this Committee might have discussed before this is relevant as well. It could be that that formula that you use in order to progress the

population year-on-year to get your estimate—predominantly inward migration and outward migration—is wrong. There are quite a lot of factors. In Manchester, partly because of our experience, on top of the census, we have our own forecasting model that we use, but again, it uses, as its foundation, the census. It just adds more in, takes more out in a different way from the way the ONS did.

Q130 Chair: This raises a question. This question touches on the whole quality of what Government has been doing for a number of years around this whole issue. The Newham surprise should not have been a surprise. The Treasury Committee remarked on how unreliable the International Passenger Survey data was for forecasting local populations. We knew that already. How effectively is the ONS handling this whole issue?

Sarah Henry: For urban areas, it could be more effective.

Q131 Chair: What should we be recommending in our report that the ONS should be doing that they are not doing?

Piers Elias: We need a bigger survey than the International Passenger Survey to catch people who are actually going to stay here every year.

Q132 Chair: We know that on that particular issue. I am talking generally about conducting this consultation, preparing for more administrative data. What are the top three things that the ONS should be getting on with that they are not getting on with?

Keith Dugmore: Better access to administrative data really is the key here. You have mentioned international migration. I understand there is the e-Borders system that, in principle, tracks people going in and out of the country, but I do not think it is used for statistical purposes

Q133 Chair: Since we produced our report on migration statistics, the Government are pressing very hard on e-Borders information, but that is something that one Department is doing in one particular area. What should the ONS be doing in this general area of understanding our population and their habits?

Sarah Henry: They should make sure that they do not disaggregate the census from the work that is done in between censuses. It is important to look at the whole package. The other thing they should be doing is exercises that match data based on identifiable information. The fact that they are trialling anonymised data, in my opinion, is flawed and unnecessary, because the census itself is not anonymised. That is very much identifiable, so stick with that.

Piers Elias: The ONS should be lobbying you to introduce some sort of compulsory re-registration when you move for local government services, or offer some incentive to make more people sign up to their GP, because the GP patient register is crucial to ONS estimates between the censuses, or to make sure the electoral register is 100%. 10% of the country is not on these registers.

Professor Mayhew: They need to get their skates on. They ought to carry out a lot more local studies in the vein that we have been testing in the past. They should then take that learning and feed it into the system that is going forward from 2020. These open-data

issues are incredibly important. One of the things that happened with the passing of the Health and Social Care Act in April 2013 was almost like data coup d'état. All data that were previously available at primary care trust level, with their abolition, migrated into the Health & Social Care Information Centre. No personally-identifiable or address-identifiable data is allowed to come out of that system.

This is really the left hand not knowing what the right hand is doing, because, in the meantime, we have a public health function moving to local authorities, the setting up of health and well-being boards, which rely on an accurate and up-to-date population to conduct any number of strategic and local issues. All that data, now transferred to the centre, is denied to them, and all that data is data they previously had access to. They are the very people who have the capacity to link that data and use it. It is a totally bizarre situation, and I do not know why it has not had more publicity than it has had, because that is driving a wedge.

It does not apply across the whole United Kingdom. In Scotland, where I also do work and I am helping them, that is not the case with these data. It really is a case of the left hand not knowing what the right hand is doing.

Q134 Chair: Are there any other comments?

Keith Dugmore: Just to pick up further on administrative data and something Sarah said, it is important that the census is not seen as a single object. The whole process of producing neighbourhood statistics and more open data is really important. I would encourage the ONS to start producing aggregate statistics from more administrative files and publishing them, albeit saying that they are experimental statistics or whatever. Culturally, there is often a feeling that we cannot publish it until we know it is perfect. If data were put into the wider domain, I think that users would be able to say, "Well, this looks pretty good" or "There are some difficulties here." To begin to churn out statistics that may be useful to people, while they cannot be perfect, would help us get towards better data for the next census.

Piers Elias: Starting with the 2011 census, the ONS published alongside the census estimates plausibility numbers from different administrative data sources, and they do intend to start publishing them in the mid-year estimates. I am not sure they are going to do it for the 2012-13 set, but we are hoping they will. That will give local authorities more confidence in seeing that, for example, their five-to-nine year olds are within the figures that the pupil censuses are saying. We would encourage them to do that.

Professor Mayhew: On that point, if you use administrative data, you can get the exact age of individuals, whether they are 5.5 years old or whatever, so you are not forced into artificial age bandwidths for the purposes of developing a new service. Frequently, with the census, you are forced into certain age groupings, which may or may not be appropriate for the service you are trying to calibrate and provide.

Q135 Chair: This is just something that is clarifying in my mind. There are three kinds of data. There is census data, which attempts to collect from 100% of the population. There is administrative data, which attempts to collect from 100% of the population. Then there is survey data, which is the sampling so championed by my colleague, Mr Flynn. The

administrative data, alongside the surveys, could replicate a great deal of what is in the census. The question is how much, and whether we have the capability to do that. Have I summarised it correctly?

Sarah Henry: In terms of the counts, that is right. There is still a question about people who are not registered on any of the administrative datasets for good reason, or counted on administrative datasets and do not match the criteria we would need for censuses. For example, people living abroad, and that kind of thing, are still registered with GPs locally, or can be.

Q136 Chair: Those tend to be the people that a census misses anyway.

Sarah Henry: It depends what the weather was like when the census was conducted. Our big issue is with the proposals around the surveys now, and the information that we would get at a level that matters to our local councillors—ward levels and even lower than that. At the moment, we think that the administrative data solution needs developing further, but it can be good for the count, and maybe the general demographics. In terms of what the survey would add, we think it is a long way from being a good solution. We need to think about how to boost it, how to change it and how maybe certain areas, like urban areas, could add to it, but it still needs a lot of development in the way it is presented in the consultation.

Q137 Chair: My last question, in summary, is how many marks out of 10 would you give the ONS for the conduct and supervision of population data?

Sarah Henry: Six.

Professor Mayhew: Is this off the record?

Chair: I guess that is less than 6.

Professor Mayhew: Well, I think they do some things well and other things not so well.

Chair: Average score?

Professor Mayhew: Five or 6—that is an average.

Kelvin Hopkins: This is just a resource problem, though. If we got the immigration statistics right, we would solve a lot of problems.

Juliet Whitworth: I have to do an average of my local authority colleagues, so I would say 6.5.

Keith Dugmore: I would go for a mean of 7.

Piers Elias: Seven.

Chair: Thank you very much indeed. It has been a very informative session. I hope you feel you have communicated what you needed to tell us. Do send in any further comments if they occur to you, and we can publish those as evidence or just informally to help us draft our report. Thank you very much indeed.

