

CONSTITUTION, EUROPE, EXTERNAL AFFAIRS AND CULTURE COMMITTEE

20<sup>th</sup> Meeting, 2022, Session 6

22 September 2022

Scotland's Census

1. The Committee has previously taken evidence from National Records of Scotland (on [23 June](#)) and representatives of the International Steering Group of global census experts established by the Registrar General (on 8 September, Official Report still to be published online, but the text is included at Annexe C).
2. At this meeting, the Committee will be taking evidence from the Cabinet Secretary for the Constitution, External Affairs and Culture and the Director of Culture and Major Events, Scottish Government; as well as the Registrar General and the Director of Statistical Services, National Records of Scotland (NRS).
3. The following papers are attached—
  - **Annexe A:** Letter from Mark Pont of the Office for Statistics Regulation to Pete Whitehouse of NRS, 17 August 2022
  - **Annexe B:** International Steering Group – Key Points, 25 August 2022
  - **Annexe C:** Official Report of CEEAC Committee meeting on 8 September, 2022
4. Members may also find it helpful to refer back to briefings provided for the [8 September](#) meeting with members of the Steering Group, and the previous session with NRS on [23 June](#).
5. A letter from NRS to OSR, a response to the correspondence included at Annexe A, is still to be published but a link will be provided to the Committee in advance of the meeting on 22 September.

CEEAC Committee Clerks  
September 2022

# Mark Pont to Pete Whitehouse: Scottish Population Census

Dear Pete

## Scottish Population Census

The Census in Scotland continues to be a major undertaking by National Records of Scotland (NRS). We welcome the ongoing engagement between your team and ours as we carry out our assessment against the [Code of Practice for Statistics](#). I am writing now, ahead of our formal assessment report which we hope to publish in the autumn, to outline some key expectations of NRS that will form a crucial part of our assessment of compliance with the Code of Practice.

The disruption caused by the COVID-19 pandemic and the change in both timing and mode of data collection to [digital first](#) mean that the context of this census is noticeably different from previous ones. There has been much coverage recently of the level of response and concerns raised about the possible impact of this on the quality of population estimates from the census. We share your concerns that the focus on the census response rate as the sole indicator of the quality of estimates could become a distraction. In particular, we consider that questions about what response rate is sufficient do not adequately capture that population outputs do not rely on census data alone, but a combination of census and other data.

Having said that, we welcome the steps taken by NRS to achieve the best response rate possible and to maximise the usefulness of the census coverage survey. The additional steps that NRS is taking, such as investigating what additional non-census data can be used to further improve the quality of census estimates and establishing an International Steering Group to provide additional expert input, all point towards an endeavour that continues the aim of producing the best census outputs possible.

In the lead up to census day, we saw NRS be open and transparent about its planned approach to [quality assurance](#) and use of [administrative](#)

[data](#) through publishing materials on its website. We also note your commitment to provide detailed supporting materials alongside census outputs to help users make best use of census estimates.

However, given the range of challenge that has been evident over recent months, we consider that it would be in NRS's interests to be more transparent now about the steps that it is taking to generate good quality census estimates. We consider that being transparent about the various current activities, plans, processes etc would assure users of NRS's trustworthiness and reassure users that they can confidently expect high quality estimates from the 2022 census. Various practices within the Code of Practice<sup>1</sup> reinforce this need for openness even ahead of the release of the statistics. Demonstrating this trustworthiness is especially vital given the importance of census estimates.

It will therefore be important for NRS to update the existing materials soon to reflect how you are adapting your processes, and the possible impact on the expected quality of census estimates. We are mindful of the need to be proportionate in doing this, especially when the focus of your team is on processing census data. At the very least we expect a clear description of how the various steps you are taking help to assure the quality of census estimates fit together, and to provide updates to users about the level of quality that they may expect from census outputs.

Thank you again to you and your team for engaging with us through this process.

Yours sincerely

Mark Pont  
Assessment Programme Lead

**International Steering Group– Note of Key Points**

25th August 2022 : 9am – 11 am (BST time)

**Attending**

Professor James Brown – UTS  
Abby Morgan – Stats New Zealand  
Owen Abbott – ONS Professor  
David Martin – University of Southampton  
Sir Ian Diamond – National Statistician  
Professor Andrew Morris - Health Data Research UK  
Alasdair McAlpine – SG  
Caroline Ellis – NRS  
David Rowley – NRS  
Benjamin Little – NRS  
Ralph McDevitt – NRS  
Lyndsay Wilson – NRS  
Esta Clark – NRS  
Tom Macintyre – NRS  
Jenny Allan-Jones – NRS  
Will Howes – NRS  
Mark Dunn - NRS

**Key Actions**

1. The group discussed attendance of ISG members at the Constitution, Europe, External Affairs and Culture Committee meeting on 8 September 2022. It was agreed that some members of the group will attend to give evidence relating to the group's role.
2. NRS provided an update on the near-final results from the CCS – An interview was completed for 58% of households in the CCS sample, while 22% provided a hard refusal and a self-completion form was left with 19%. The lowest interview rate was 52% in COMA 7, while the highest was 67% in COMA 4. The numbers may change slightly as some self-completion forms are still being received. NRS will continue to accept these until around mid-September.
3. The group provided feedback that the extension to the CCS was valuable, enabling all COMA areas to achieve at least a 50% interview rate.
4. NRS provided an update on work to match CCS and Census addresses, providing results split by Hard to Count area. NRS will provide an update on vacant property rates in the CCS and Census at a future meeting.
5. NRS provided an update on the application being prepared for the Health and Social Care PBPP for the use of administrative data in the census. This included discussion of whether to apply for permission to place admin data records directly in the census and then keep them in the final census dataset. Potential benefits of this would be a higher number of records based on genuine individuals included in the

census microdata, which is made available to researchers, and in the release of census records in 100 years.

6. The group provided a strong steer that NRS should focus on the main purpose of using admin data to enhance the estimation and adjustment processes and improve the accuracy of core census outputs. The application should emphasise the key benefits of using admin data to improve the accuracy of both the high-level and lowlevel population estimates. This does not require admin data records to be placed directly in the final census dataset, so the group advised not to include this in the PBPP application.

7. NRS agreed to consider this advice and circulate a revised PBPP application to the group for comment.

25th August 2022



OFFICIAL REPORT  
AITHISG OIFIGEIL

DRAFT

# Constitution, Europe, External Affairs and Culture Committee

Thursday 8 September 2022

Session 6



The Scottish Parliament  
Pàrlamaid na h-Alba

© Parliamentary copyright. Scottish Parliamentary Corporate Body

Information on the Scottish Parliament's copyright policy can be found on the website - [www.parliament.scot](http://www.parliament.scot) or by contacting Public Information on 0131 348 5000

---

**Thursday 8 September 2022**

**CONTENTS**

	<b>Col.</b>
<b>DECISION ON TAKING BUSINESS IN PRIVATE .....</b>	<b>1</b>
<b>SCOTLAND'S CENSUS .....</b>	<b>2</b>

---

**CONSTITUTION, EUROPE, EXTERNAL AFFAIRS AND CULTURE COMMITTEE**  
**19<sup>th</sup> Meeting 2022, Session 6**

**CONVENER**

\*Clare Adamson (Motherwell and Wishaw) (SNP)

**DEPUTY CONVENER**

\*Donald Cameron (Highlands and Islands) (Con)

**COMMITTEE MEMBERS**

\*Alasdair Allan (Na h-Eileanan an Iar) (SNP)

\*Sarah Boyack (Lothian) (Lab)

\*Maurice Golden (North East Scotland) (Con)

\*Jenni Minto (Argyll and Bute) (SNP)

\*Mark Ruskell (Mid Scotland and Fife) (Green)

\*attended

**THE FOLLOWING ALSO PARTICIPATED:**

Professor Sir Ian Diamond (UK Statistics Authority)

Professor David Martin (UK Data Service)

**CLERK TO THE COMMITTEE**

James Johnston

**LOCATION**

The James Clerk Maxwell Room (CR4)

## Scottish Parliament

### Constitution, Europe, External Affairs and Culture Committee

*Thursday 8 September 2022*

*[The Convener opened the meeting at 09:00]*

### Decision on Taking Business in Private

**The Convener (Clare Adamson):** Good morning. I give a warm welcome to the Constitution, Europe, External Affairs and Culture Committee's 19th meeting in 2022. Under our first agenda item, do members agree to take item 3 in private?

**Members** *indicated agreement.*

## Scotland's Census

09:00

**The Convener:** Our second agenda item continues our work on Scotland's census. We are joined remotely by two members of the National Records of Scotland international steering group. I warmly welcome Professor Sir Ian Diamond, the national statistician, from the UK Statistics Authority, and Professor David Martin, a professor of geography at the University of Southampton and deputy director of the UK Data Service. Good morning to you both.

I will begin with a couple of questions. When the Office for Statistics Regulation wrote to the National Records of Scotland on 17 August, it said:

"The disruption caused by the COVID-19 pandemic and the change in both timing and mode of data collection to digital first mean that the context of this census is noticeably different from previous ones."

That was not covered in your group's response to the committee, which we thank you for, but you are very close to the topic. Will you elaborate on the context for the census and its differences from previous ones? I invite Sir Ian Diamond to go first. *[Interruption.]* We cannot hear you.

**Professor Sir Ian Diamond (UK Statistics Authority):** Can you hear me now?

**The Convener:** We are fine now.

**Professor Sir Ian Diamond:** Good morning—I am glad that you can hear me. In effect, you have asked two questions; I will take the second one first and the first one second. Following the United Kingdom Parliament's decision in 2014 to undertake a 2021 census, it was agreed that, given improvements in technology and its accessibility, the 2021 census would be the first digital-first census. The Office for National Statistics put an enormous amount of work into ensuring that it was possible for citizens to fill in their census form entirely digitally and that the forms were digital friendly, so that they could be filled in easily on a mobile device as well as on a more traditional computer. Such work was successful.

It was entirely recognised throughout that some citizens do not have access to digital means, so paper questionnaires were produced. In areas where it was expected that, for example, broadband use or availability would be low, paper questionnaires were sent out en masse. Elsewhere, the approach was digital first, and paper questionnaires were used only when enumerators went to pick up places where digital responses had not been given, so that enumerators could say, "I have a paper questionnaire."

That was a difference, but it was different only in as much as the way of filling in the questionnaire

differed. Much of the methodology—such as the definition of usual residence and the individual-level census—was as it has been since 1841, but with relevant questions for 2021.

You mentioned that, in 2021, the censuses of England and Wales and of Northern Ireland took place when there had been a pandemic. As I am sure that you are aware, that was not the first time that that has happened. In 1921 there was a short delay to the census because of a big wave of flu. The context was not new, therefore—the same thing has happened before, albeit some time ago. We at the ONS and my colleagues at the Northern Ireland Statistics and Research Agency have looked very carefully at all the preparation and all the considerations, and we took a judgment that we would undertake a census in 2021. I have to say, by way of finishing, that it was an unbelievably successful census.

**Professor David Martin (UK Data Service):** Sir Ian has outlined two major changes for you. If you are talking about the broader context of the census, there a couple of other factors.

You are looking as if you might not be able to hear me. Can you hear me okay?

**The Convener:** I am struggling slightly. Can we have the volume turned up, if possible? We might be able to fix the issue in the room. Please continue: it is probably just me.

**Professor Martin:** The two lesser factors that I would add in the broader context of the census are a general societal change towards a lower response to censuses or surveys, which we see year on year. The continuous level of survey response has had a gentle, steady decline, and the circumstances of the pandemic will have meant that, certainly in England and Wales, lives were disrupted. Working patterns were quite different, with many more people working from home, and the ways in which people felt that they should correctly answer questions, given the contextual factors, would have been quite different by the time of getting to the census. Ian Diamond has given you the most important trends of the principles, and I completely agree with his comments.

**The Convener:** As I have asked two questions in my first question, I will now move to other members of the committee, starting with Mr Cameron.

**Donald Cameron (Highlands and Islands) (Con):** Sir Ian, you spoke about what you saw as a very successful census in the rest of the UK in 2021. I think that the completion rate was about 97 per cent. The Scotland census reached a figure of 87.9 per cent—9 per cent lower. In Glasgow, our biggest city, we got only as high as 81 per cent or thereabouts. Why do you think that that happened? Why is there a disparity?

**Professor Sir Ian Diamond:** Thank you for the question. I apologise, as I must simply say that I am not really able to answer that, for the simple reason that it is entirely necessary and appropriate—we did this in England and Wales, as did my colleagues in Northern Ireland—to undertake a lessons-learned exercise and an investigation of what went well and what did not go so well. I am not aware that that work has been done yet, and I have not been asked to review such work. Anything that I could say would simply be speculation, and I do not think that it would be right to speculate at the moment.

My view has always been that, given the urgency, the important thing was to move forward and to get to a position of having top-class population statistics for Scotland available in the spring of next year. It is important that an assessment as you have described is made. I would be delighted to be part of that if asked, but that is for the Scottish Government.

**Donald Cameron:** We are hearing from the Scottish Government next week, so that is something that we could easily take up with it.

I also ask Professor Martin for his view, if he is able to comment—I appreciate that you may not wish to.

**Professor Martin:** Again, I emphasise that the international steering group very much sees Scotland's census as an operation that is still taking place. It is not finished when the count is over. It is not part of our remit to investigate how you got to where we are at this moment; we are endeavouring to advise the NRS on the best steps to take right now. That investigation needs to take place and has not yet happened.

I will add one point, as Donald Cameron picked up on response rates in, for example, Glasgow. When we look at the census response in recent censuses across the whole of the UK, there is, in fact, considerable variation between local authorities. Some large local authorities in England and Wales are also more challenging to enumerate, and we see a range of response rates. Although it is disappointing that what we are seeing from Glasgow is not higher, it is not in a completely different ball park to figures that we have seen from other large cities in recent censuses. However, delving into the exact reasons as to how you got there is yet to be done.

**Donald Cameron:** I will turn to a slightly different question, which I hope that you may be able to help with. Looking forward, on what statistics in particular do you see there being an impact from the difference between the rest of the UK and Scotland?

**Professor Sir Ian Diamond:** If I may, I will describe where I hope that we can get to. As

Professor Martin has indicated, these days, a census is not simply the initial data collection exercise. Certainly, that initial data collection exercise is the most important pillar, but three pillars are brought together that make a population statistics system.

The first pillar is the census. The second pillar is a coverage survey, in which we go back to a sample of postcodes, redo the census and then link those together to make estimates of the underenumeration. As Professor Martin has pointed out, we expect beforehand that there will be higher underenumeration in those areas in Scotland that would be lower on the Scottish index of multiple deprivation, and we design for that. After the coverage survey has been linked—we use dual-system estimation to make estimates of the missing people and households—we then use administrative data to do quality assessment and further imputation. That is the third pillar.

In relation to the first pillar, the response is a little lower than we would have hoped for. We have done the coverage survey, and the steering group has been giving the NRS close advice on all matters relating to that. Indeed, when there was difficulty with responses in some areas, I took a judgment that the ONS would pause some of our data collection for other purposes in Scotland and some of our professional interviewers went to help. We have been doing everything that we can to get that good.

We still have the administrative data to do, which it will also be necessary to bring in, and for which we use different methods. There are proposals on using quite innovative statistics around administrative data to make estimates. As I said earlier, when we have those three pillars together, I very much hope that the NRS will have some really reliable population statistics for Scotland by the spring of the next calendar year.

At the same time, we in England and Wales, and my colleagues in NISRA, are working very hard on moving our 2021 estimates through to 2022. We are using a dynamic population model. That new method is extremely accurate and based largely on administrative data. We will then have estimates for 2022—all with confidence intervals and statistically sound—which will be directly comparable with those in Scotland.

09:15

In summary, I am expecting—I am hoping—that we will be able to have directly comparable UK-wide population statistics for 2021 by spring next year. I do not say that with complete confidence, as the work on the third pillar—the administrative data—still needs to be done. A further problem that we will need to look at is that the initial starting point for some of the statistics that we do assumes

independence between the census and the coverage survey. We expect there to be some dependence, and we will have to estimate that.

There is a lot of quite complex statistical analysis still to do in Scotland. The ability to access very good administrative data is absolutely critical. Given those two things—I have confidence in the statistical analysis—I believe that it will be possible to have UK-wide population data by the spring of 2023.

**Donald Cameron:** Thank you for that. David Martin, do you want to add anything?

**Professor Martin:** To assist the committee, I will mention that there is a broader international context of shifting the emphasis between the three pillars. You will have read the report. Ian has emphasised the importance of the initial enumeration, which is followed by the census coverage survey and the work with the administrative data. We are seeing a journey on which there is increasing reliance on the administrative data sources in producing the complete national population estimates. Therefore, the committee is capturing the situation with Scotland's census precisely as that transition is taking place. There will be administrative data in the Northern Ireland census results from 2021, which are just being published. In Northern Ireland, administrative data have been used to complete the record.

We are talking about known waters internationally, but the circumstances are slightly different in each of the UK countries when we look at how precisely the census balance is working out—hence the emphasis that is now being placed on the administrative data.

**The Convener:** Dr Allan has a supplementary.

**Alasdair Allan (Na h-Eileanan an Iar) (SNP):** You have both alluded to the fact that people in many communities in large cities are increasingly resistant to filling in surveys. Will you explain what you mean? Can you suggest any reasons for that?

Perhaps Sir Ian could go first.

**Professor Sir Ian Diamond:** It is lovely to see you, Dr Allan. I hope that you are keeping well; I have not seen you for a while.

There is a host of reasons why survey response rates have been going down over the past few years. First, the fact that people are not at home so often or so regularly means that it is difficult to contact them. Secondly, people do not always see it as their civic duty, if you like, to help with surveys. Thirdly, there are so many surveys that people are never quite sure about their importance. We make enormous efforts to impress on people the fact that the surveys that we are talking about are important Government surveys and to explain what they are used for. When we make the distinction between

**!Syntax Error, !**

**ERROR! UNKNOWN DOCUMENT PROPERTY NAME.**

**!Syntax Error, !**

such important Government surveys and other kinds of survey, people tend to be more likely to help.

There have been reductions in survey response rates across the world. An enormous amount of research is being done across the world on how to improve response rates. Indeed, just this week, the UK Economic and Social Research Council announced a big programme of research on improving response rates. The Office for National Statistics will provide support in kind to researchers who will use some of our surveys to do experiments on how to improve response rates.

A multifarious group of factors is at play, and linked to that is that people, who are very busy, feeling that there are lots of surveys and wondering why should they answer them. It also to do with people not being there when you knock on the door, as well as, perhaps, a reduction in civic duty.

**The Convener:** Does Professor Martin want to comment?

**Professor Martin:** Sir Ian has given you the principal dimensions of the matter. As a geographer, one of the things that I have been very much involved in over time is the process of address listing and looking at the way in which data work for small areas.

There is an additional factor: increasingly, people live in properties that are hard to access. There is an entry system that is remote from the individual's front door and it becomes increasingly difficult to deliver direct to the door for surveyors and census coverage survey surveyors to gain access to those individual people to work out who is at home. That situation is, of course, traditionally prevalent in large cities.

There are structural, address-related factors to the way in which people's housing is arranged that additionally drive the situation over and above the social component that Sir Ian has just highlighted.

**Jenni Minto (Argyll and Bute) (SNP):** I thank the witnesses for joining us. I will ask a similar question to the one that I asked Mr Lowe of the National Records of Scotland. The international steering group's letter noted that the census results provided a strong foundation. Will you explain to me, a layperson, what that means?

**Professor Martin:** I go back to a point of emphasis: you are still conducting your census and it is a three-pillar exercise. The ideal would be that you conducted a census and needed only the first pillar because you lived in a fairyland where everything was done perfectly, everybody was compliant and every form was returned. That is not the world that we live in. The level of response that you have, which is slightly short of 90 per cent—in the upper 80s overall—is sufficient for us to be able

to build good estimates if the other pillars provide what is needed. There is always a cross-reliance on those.

You will not be surprised that the steering group thought hard about those words and is confident that getting an 89 per cent response is a good foundation for doing the remainder of the work. That decision was taken in the full knowledge that a lot of ground work would be needed on the administrative data and without knowing at that point what kind of success rate we would see from the coverage survey.

I do not think that any of us would want to revisit that. We are quite content that that response rate allows you to make good estimates but it is dependent on the other parts coming in and performing in a way that you can stitch together to get the whole.

**Jenni Minto:** Professor Diamond, do you have anything to add?

**Professor Sir Ian Diamond:** I do not really have anything to add but, as I said previously, the quality of the administrative data will be critical and you should be assured of two things. First, the statistical analysis that needs to be done is hard and we are lucky that the UK has developed some of the techniques so we have the experts who are able to help throughout the UK. I will ensure that the Office for National Statistics is able to provide any support that is necessary. Secondly, we are also privileged that Professor Brown, who is an international expert, is chairing the steering group and will provide advice on some pretty hard statistical analysis.

**Jenni Minto:** You made a point, Professor Diamond, that, in 1921, the flu pandemic delayed the census for a while. Do you have any thoughts on what happened? Were there delays in other countries as a result of the pandemic?

**Professor Sir Ian Diamond:** The picture is mixed. Some countries undertook their census at the same time—indeed, some of them did theirs in 2020—and others, such as the Republic of Ireland, delayed by one year, so no algorithm exists that enables me to say, “this is what happened”.

I am conscious that every country did as I did with my colleagues in the ONS and considered all kinds of indicators, undertook all kinds of preparation, talked to many people—for example, we took advice from the chief medical officer about some issues—and came to a judgment. Again, I did not have an algorithm or a computer that told me the answer. We came to a judgment and made a recommendation to our board, which our board accepted, and I believed it to be the right recommendation for England and Wales.

**Jenni Minto:** I suppose that some of your considerations would have involved the availability of the resources that you have as an organisation and the other work that you were carrying out at that time. I believe that the structure is slightly different in Scotland, with the NRS office being much more involved in recording Covid deaths.

**Professor Sir Ian Diamond:** Our job is also to record Covid deaths. Our weekly production of mortality statistics during the pandemic, a sad task, was a major source of information for the national health service, the Department of Health and Social Care and the Government in England, on which they could make judgments. I had a big team working on mortality statistics, which I have to say did a fabulous job in very difficult circumstances.

Throughout the pandemic, we were also producing gross domestic product figures and inflation figures, which are important at the moment. Some of those things had to be changed at real pace—if you cannot go into a shop to collect price data, how do you calculate inflation data?—so an enormous amount of extra work was done there.

We were considering all those factors as well as preparing for a census. As an aside, you might see the weekly Covid infection survey, which we design, because Scotland is part of those results. That survey was set up from scratch—150,000 swabs taken every fortnight and analysed in a complex way—to produce weekly statistics.

We were not just an organisation doing a census. An awful lot of things were going on and we needed to keep all those things going on—the nation could not exist without inflation figures or without knowing what the labour force is doing. We needed statistics on weekly Covid deaths and the infection survey. We needed all those things, and the judgement was that we could also do a very good census—and we could.

**Jenni Minto:** Thank you for that, Professor Diamond. I agree that the work that you do is incredibly important, as is the work of the NRS office.

**Professor Sir Ian Diamond:** I was in no way suggesting that the work of the NRS is anything other than incredibly important.

**Sarah Boyack (Lothian) (Lab):** Thank you, Professor Diamond, for the briefing that you sent us, which has been useful. I have follow-up questions about the post-census work to which you refer in the briefing.

How do you fill the gaps that come from the higher non-response rate than we had in 2011? How do you avoid errors in the assumptions that are made in the final stage—pillar 3, which you talk about—in order to add value to the census returns

that we already have? How do you ensure that the information that you add to the census will give confidence to people who use it—particularly in the lower response rate areas, to which you refer? How do you make that calculation about the geography of those areas and the different groups of people who have not filled in the census? How do you avoid errors there? What assumptions are made and how do you make sure that there is no bias in those assumptions? You talked about that being useful in relation to what groups might have been excluded.

Professor Martin, you talked about the difficulty of access to buildings. There are also buildings that are easy to access but produce incredibly low results. What is your perspective on how to get that right for the people who will rely on the data in the census?

09:30

Professor Martin, do you want to kick off, because you mentioned the issue of access? I live in a city that has loads of tenement flats, so there are always access issues. In the big place that I visited with the enumerators, what struck me was not the access but the massively low turnout; it was less than 50 per cent, and that was in the boost period after the census had officially finished.

**Professor Martin:** You raise important issues that are well recognised in the statistical agencies, because they are inherent to the census process. It might be helpful to explain a little more about the way that the coverage survey works and how the administration data are brought into play.

Before the census is conducted, the NRS will have conducted an exercise that, in effect, classifies areas according to how hard they will be to count—it is actually called a hard-to-count index. It is based on what we know to be the drivers from census responses in the past. We know the various kinds of relevant factors, such as the sorts of housing, the concentration of students and language obstacles—we know all of the characteristics that come together to make it difficult to get responses. That is quite strongly aligned with the SIMD, but it is not aligned only with deprivation.

That information is used to carefully stratify the areas that are targeted in the census coverage survey in pillar 2, so that much more emphasis is put on going back to those areas where we know census response will be difficult. There are five bands in the coverage survey, and, as an indication of the emphasis, the easiest 40 per cent of the country goes into band 1, whereas the hardest 2 per cent goes into band 5, which includes the areas that are given that extra emphasis. Twenty times more effort is put in in terms of the intensity of

surveying those more difficult areas. In Scotland's case, that is spread across all the council areas.

The coverage survey—which is only a small sample—aims to go back to find where we have under-response. Can we find the same people who responded to the census when we go back again? That gives us a good picture of the way that different neighbourhood types in different council areas have responded, and gives us a detailed matrix of the under-response.

If the coverage survey had very high coverage, that would allow us to do the estimation and receive the correction. What we are looking at here is that that process will tell us that we need to bring in the administrative data, because the administrative data often tells us about the presence of addresses from which there has been no response, even though there is plenty of administrative evidence that there are individuals living at those addresses. That would then feed in to the estimation of the total numbers before the adjustment is made.

The CCS and the administrative data are very targeted on precisely that question and designing the system to avoid bias and not just fill in the people who are easy to find, which is the point that you rightly raise. That is central to the way that the system is designed from the start.

**Sarah Boyack:** Professor Diamond, do you have any comment on the issue of how to avoid bias in low-response areas?

**Professor Sir Ian Diamond:** First, I agree with everything that Professor Martin just said, but I will add a couple of points if I may.

In many ways, your point is really about bias. Bias can come about in the statistical analysis because people who do not respond to the census are also more likely to not respond to the coverage survey. We call that “dependence”, and that can lead to bias in the analysis that leads to your estimates. We recognise that. Indeed, along with Professor Brown, who chairs the committee, and Owen Abbott, who is also on the steering group, I wrote a paper about the issue in 2006. One way in which the administrative data can come in, in the way that Professor Martin has just described, is by helping us to estimate that dependence and then to adjust for it. That is incredibly important.

The second thing that it is important to do, which we have not yet talked about, relates to communal establishments. The administrative data are really important in relation to those, and that is particularly relevant to student halls of residence and care homes. It is important to get good administrative data for those so that biases do not come about through underestimating populations in communal establishments. They represent only a small proportion of the population, but if you do not get

them, an important part of the population will be missed.

**Sarah Boyack:** From looking at previous census data, do you have a sense of the differences between the 2022 census and the previous census in 2011?

**Professor Sir Ian Diamond:** I have said clearly that I am pretty sure that we can get some really great estimates, but one of the challenges that statisticians face is the design of the survey. If you were to ask me to design a survey to estimate the proportion of something, my first question would be, “What kind of proportion do you think that it might be?”, because the sample size and the overall design will be based on that estimate. The aim in Scotland was a response rate of 94 per cent. Therefore, initially, the design was for an underenumeration of 6 per cent. As the underenumeration increases, your confidence interval potentially increases. As I said to someone the other day, as you get further out—we are not too far out—the confidence intervals, as I am demonstrating with my hands, do not increase slightly but dramatically. When you get further out than we are in Scotland, they become as wide as an albatross’s wingspan.

Therefore, you do not get bias—I have talked about bias—but you must be aware that you will lose precision. However, in Scotland, although we are in a position where, yes, we are going further out than we want to, and, yes, we have less precision than we had estimated, we are still in a position, if we can control for all the biases, to make estimates at a level that has been achieved elsewhere in the world. Indeed, I can think of some places that have potentially had bigger problems than those that Scotland is facing at the moment. The situation is not impossible; it is hard. Scotland now has some of the best people in the world advising it, and as long as the administrative data are good, we can control for bias, maximise the precision of the estimates and get to a place where you have really good, useful population estimates that are comparable with those across the UK.

**Sarah Boyack:** That is great. Will the administrative data be published separately, or is that integrated into the final census results, so that it is transparent?

**Professor Sir Ian Diamond:** It is integrated, because there are clearly ethics and privacy issues to be taken into account. Transparency is a matter for the NRS, but I would expect that transparency lies in a clear exposition of the methodology, not in the publication of data, which would impact on ethics and privacy.

**Maurice Golden (North East Scotland) (Con):** The written evidence that we have received on international approaches to census taking notes

that the use of administrative data, as you have highlighted today, is ideally part of a process for quality assurance on the final census output. However, you have said that the use of administrative data in Scotland's 2022 census will be central to the final quality of that census. As we have heard, that is necessary as a result of the low response rates for the census and the community coverage survey, which both missed their targets. Does relying on the use of administrative data in that way fall short of international best practice? Perhaps Professor Ian Diamond can start.

**Professor Sir Ian Diamond:** I do not think that it does. Both Professor Martin and I have, on a number of occasions this morning, pointed to the fact that we are still doing a census in Scotland. The phrase "quality assurance" can mean making further adjustments. For example, in the census of England and Wales in 1991, administrative data were used to identify that the census had missed quite a large number of men aged between 20 and 34. Using those administrative data as a base, an adjustment was made to add in more men, and the final results followed. That was not just about quality assurance. It started with quality assurance, but a problem was then identified, and it was solved by further adjustment. The overall numbers were then based on administrative data and, in that case, on the census, because we did not have a coverage survey in those days.

That approach would not fall away from international best practice. There are some suggestions about bringing some of the administrative data into the coverage survey. That would be innovative and exciting, but it would not be against international best practice.

Professor Martin will probably have a view as well.

**Professor Martin:** I can illustrate the situation for you. I recall a presentation that I gave in 2017 to a group of young social scientists at a summer school or something of that kind. I was invited to talk about the international situation, looking forward to the 2021-22 round of censuses. I produced a slide that plotted different countries and demonstrated that all of them were moving along a trajectory from a reliance on conventional enumeration and nothing else towards more and more use of administrative data.

As we have seen over the past two or three decades, plenty of countries have shifted completely from a conventional census and enumeration to producing their population statistics entirely from administrative data, augmented by some surveys. In the UK, with the three UK censuses, we are somewhere in the middle. We have a hybrid model in which we use administrative data at the aggregate level—for example, we find out how many schoolchildren the school census

shows should be in a small area and how well that matches with what we have from the census record—and, increasingly, to fill gaps, to use the committee's term.

I will give one example that has not been mentioned. Surprisingly, people often leave young babies, especially those under one, off census forms. People do not necessarily read the instructions on the internet; that is a recognised phenomenon. We do not ignore that fact—we routinely take a look at birth registration data to work out where those people are missing.

09:45

A couple of decades ago, that might have been seen as simply a quality assurance process. However, now, if we know that those births are missing, the view is that they should be used, either directly or indirectly, to adjust the estimate and that we should fill in those that are missing. Increasingly, the same would be the case for students and, as Sir Ian Diamond has mentioned, people in communal establishments.

Therefore, the use of the administrative data in various ways is a transition from quality assurance—which says, "This has worked well, but there's an individual group that needs adjustment"—to seeing the administrative data as part of the whole design, which is characteristic of what we see going on in many countries. There are lots of axes, so it is not a simple linear process. However, the general trend towards increasing the use of administrative data is clearly an international phenomenon.

**Maurice Golden:** Thank you. That is very helpful. Professor Martin, in an international context, is there any evidence that inclusion of what might be deemed controversial questions in the census has an impact on participation and completion of the survey?

**Professor Martin:** I am not sure that we could say that there is a clear body of evidence on that. Countries ask surprisingly different census questions. What would be considered acceptable in one country might not be acceptable in another, and vice versa—which is fascinating, but not always easily explained—so censuses in individual countries often contain a controversial question or two. However, I cannot think of any particular body of evidence to show that the success or otherwise of a census has been nailed to whether the wording is wrong for a certain question. I would certainly refer you to Sir Ian Diamond on that point.

The reality is that most population members tend not to be very engaged with the debate on the questions. As we have seen, they become aware of the census when the form arrives, and they largely fill it in without reading the guidance. We have not

investigated specific questions in Scotland; that is not something that the international steering group has been tasked with or has been looking at. I would not generally consider that the choice of wording for a question is one of the major drivers of the census response.

**Maurice Golden:** Thank you. That is helpful.

Professor Sir Ian Diamond, have you any thoughts on that point?

**Professor Sir Ian Diamond:** It is a very good question. Were I teaching how to do a census—as I did for about 30 years—I would say that we should be careful not to have sensitive questions. A number of questions, in particular those on income, have been tested in England and Wales. There has always been great user demand to ask about income in the census. In the 1990s and the 2000s, the Office for National Statistics did a lot of experiments to see whether doing so might have an impact on response. The evidence was mixed, and the judgment was that we should not ask about income. One of the steps that the Government has asked the Office for National Statistics to take is to produce income estimates using administrative data following the census, which we intend to do.

Following the reconvening of the Scottish Parliament, MSPs were concerned that, under the initial plans for the 2001 census for Scotland, people would not be given the opportunity to put “Scottish” as their nationality. At the last minute, the registrar general for Scotland’s office had to reprint the census forms in order to give people the option to do so, because it was felt that not having such an option might have an impact on response. However, I do not know of anywhere where a question can be seen to have definitely had an impact on response.

I am sorry, but I will add one further point. In the evidence that I have given to the committee this morning, I have indicated on two occasions that the results should be ready in the spring of 2023. While I have been sitting here, I have been informed that the current expectation is that they should be ready one year after collection finishes. I would just like to clarify that point. I was led to believe something else, but I now have clarity.

**Maurice Golden:** Thank you for that clarification and for your contributions.

**The Convener:** Thank you very much for your evidence this morning.

A few of our more subtle questions have asked what might influence people who are completing surveys. Earlier, you mentioned the number of surveys that families are receiving. Are those from local government or central Government, or are they just general ones? When might we know what policy decisions could have influenced returns? I

am thinking of the participatory budgetary surveys that now regularly go out from councils, which might not be happening elsewhere in the country.

We have lost the sound for Sir Ian Diamond again. Go ahead, please.

**Professor Sir Ian Diamond:** I will repeat the point that I was making. I do not know about you, convener, but I rarely get through a day without two or three survey questionnaires hitting my inbox. Some ask for my views on something that I might have done recently and others ask for my views on whom I might vote for in a future election. I believe that that practice is leading to a kind of survey fatigue. That is why your point is unbelievably important.

The Government needs to have impeccable public engagement so that people are seriously aware of why their opinions are being asked for and how they will be used. The Office for National Statistics has put a lot of effort into that. It is incredibly important that we have feedback so that we can say, “You said this; we did that.” That is super important, because when we talk to people, if we just ask, “Will you fill in our survey?”, they will say, “Oh, gosh—I have done this so many times.” However, if we were to say, “Look, this is a Government survey. This is what it’s going to be used for and this is how it will impact positively on your fellow citizens and your neighbourhood. It is incredibly important that we have these data”, people would say, “I will take part. No problem.” There are still issues with ensuring that we have contact, particularly with the most marginalised members of society, but the approach should be very much about ensuring that people know why they are answering.

My other point comes back to what was said right at the beginning of the meeting, about the digital census. In England and Wales, we were incredibly impressed that we had a much higher digital uptake than we were expecting—particularly among elderly people, whom we were expecting might be digitally challenged. I have no evidence for what I am about to say, but the fact that people had spent a lot of time on Zoom with their grandchildren, for example, might have meant that they were able to engage digitally. That is potentially important for the future.

**!Syntax Error, !**

**ERROR! UNKNOWN DOCUMENT PROPERTY NAME.**

**!Syntax Error, !**

**The Convener:** I am looking to my colleagues to check, but I think that we have exhausted our questions. I thank both our witnesses for their attendance.

09:54

*Meeting continued in private until 10:47.*

This is a draft *Official Report* and is subject to correction between publication and archiving, which will take place no later than 35 working days after the date of the meeting. The most up-to-date version is available here: [www.parliament.scot/officialreport](http://www.parliament.scot/officialreport)

Members and other meeting participants who wish to suggest corrections to their contributions should contact the Official Report.

Official Report  
Room T2.20  
Scottish Parliament  
Edinburgh  
EH99 1SP

Email: [official.report@parliament.scot](mailto:official.report@parliament.scot)  
Telephone: 0131 348 5447  
Fax: 0131 348 5423

The deadline for corrections to this edition is:

**Thursday 6 October 2022**

All documents are available on  
the Scottish Parliament website at:

[www.parliament.scot](http://www.parliament.scot)

Information on non-endorsed print suppliers  
is available here:

[www.parliament.scot/documents](http://www.parliament.scot/documents)

For information on the Scottish Parliament contact  
Public Information on:

Telephone: 0131 348 5000

Textphone: 0800 092 7100

Email: [sp.info@parliament.scot](mailto:sp.info@parliament.scot)



The Scottish Parliament  
Pàrlamaid na h-Alba