



HOUSE OF LORDS

Revised transcript of evidence taken before

The Select Committee on Social Mobility

Inquiry on

SOCIAL MOBILITY

Evidence Session No. 15

Heard in Public

Questions 140 - 149

WEDNESDAY 18 NOVEMBER 2015

11.35 am

Witnesses: Professor Anne Green, Professor Sandra McNally

and Dr Stefan Speckesser

USE OF THE TRANSCRIPT

This is a corrected transcript of evidence taken in public and webcast on
www.parliamentlive.tv.

Members present

Baroness Corston (Chairman)
 Baroness Blood
 Lord Farmer
 Baroness Howells of St Davids
 Earl of Kinnoull
 Baroness Morris of Yardley
 Lord Patel
 Baroness Stedman-Scott

Examination of Witnesses

Professor Anne Green, Institute for Employment Research, University of Warwick, **Professor Sandra McNally**, Director of the Centre for Vocational Education Research, LSE, and **Dr Stefan Speckesser**, Chief Economist at the Institute for Employment Studies, LSE

Q140 The Chairman: Thank you very much for coming to appear before us today; we are very grateful. As you probably appreciate, this session is open to the public. A webcast of the session goes out live and subsequently it will be accessible via the parliamentary website. A verbatim transcript will be taken of the evidence, which will also be put on the parliamentary website. A few days after the session, you will receive the transcript to check for accuracy, and we would ask that any amendments be notified to us as quickly as possible. After the session, if you want to amplify or clarify any points, you are perfectly entitled to submit supplementary written evidence to us. Before we start, would you introduce yourselves for the record?

Professor Anne Green: Hello, I am Professor Anne Green. I am from the Institute for Employment Research at the University of Warwick. For background, I work using mixed methods, so I use quantitative data sources, case studies and qualitative methods. I am currently involved in an Economic and Social Research Council-funded project on ‘Precarious Pathways to Employment for Young People?’, which hopefully provides some insights for the kinds of work your Committee is doing.

Professor Sandra McNally: I am Sandra McNally. I am a professor of economics at the University of Surrey. I am director of the Centre for Vocational Education Research at the London School of Economics, which is a BIS-funded research centre launched in March for three to five years. Stefan is also involved in our centre. We are aiming to do rigorous quantitative analysis to inform the issues of concern to your Committee about vocational

education. We are economists; we are in an economic framework, so the provision of data is absolutely essential to the quality of what we are able to do.

Dr Stefan Speckesser: I am Dr Stefan Speckesser. I am chief economist at the Institute for Employment Studies in Brighton. I have been working on youth transitions in recent years. My background is that I am an applied econometrician and have worked on labour economics and education economics for most of the last 20 years, mainly focusing on administrative data. One of the reasons why I am part of the Centre for Vocational Education Research and why, I suppose, I am here today is that I have a lot of knowledge coming from all the years of using administrative data for education outcomes and impact assessments. I have worked on a couple of projects for the Department for Business, Innovation and Skills, in particular on learning below level 2. I have recently published a series of discussion papers, practically a week ago I suppose, available on the BIS website on youth transitions into the labour market and within the labour market, focusing on long-term outcomes.

The Chairman: Thank you very much. What are the key data sets you use to analyse patterns of provision and progression into the labour market for young people, for underserved groups and the people who are called middle attainers? What are the strengths and the limitations in the data that are consequently provided?

Dr Stefan Speckesser: Generally, compared to the situation we had about five or six years ago, the data situation has improved very dramatically in that many administrative data sets in the Department for Business, Innovation and Skills were merged so that we can see who is participating in post-16 education and we can track their later employment and earnings trajectories in HMRC administrative data.

This is going to be the major resource for research on young people's transitions to the labour market and progression in the labour market in the coming years. It is a mixture of education data coming from schools, essentially, which is partially linked already, and is to be linked in the near future, so that we have an exact account about individual participation in compulsory and post-compulsory education, or post key stage 4 education in the country. That is the primary resource we are going to use.

Vocational education is a very complex field of education, with very different pathways people could choose. The available survey evidence from the large youth cohort studies—longitudinal studies of young people in England—is not really large enough in terms of size to look into particular details of education pathways and subsequent trajectories. Some of it

is available as large-scale administrative data, some of it is emerging, and we work on this as we speak. There is an interdepartmental working group preparing administrative data to be used for this purpose. The Centre for Vocational Education Research will access these data and provide research to look into education and employment trajectories.

In addition to this, traditionally, a lot of research on young people's outcomes in the labour market was based on the available cohort studies. These are people born in particular weeks, in 1958 or 1970, who are in the recent longitudinal study of young people in England, but there are various limitations when you use such data sets, one of which is that some of the cohorts are now in their 40s or 50s, so the results may no longer be indicative for today's transitions in the labour market. One problem with the cohort studies is that the longitudinal study of young people, for example, currently covers people up to the age of 20, so you do not really have the long-term outcomes in the labour market that everybody is interested in.

I can add one more thing. Available surveys, such as the Labour Force Survey, have been used a lot in the past in research on vocational education. The Labour Force Survey is the survey that is used to provide official unemployment statistics, but education is an important characteristic included in the data set, so many of the studies that emerged on the outcomes of vocational education are based on these data sets, such as returns to vocational education and the impact on adult employment trajectories. A lot of that comes from survey data, such as the Labour Force Survey, but, by and large, the level of detail that you could achieve with survey data will not be as detailed as what we can do with administrative data.

Q141 The Chairman: Professor Green, could you share with us the findings of your project on Precarious Pathways to Employment for Young People? Are there any emerging themes?

Professor Anne Green: Yes, there are certain key emerging themes. Interestingly for you, we are looking at young people who, perhaps, leave school at 16 and then go on to a college course—people who do not, perhaps, go through the five GCSEs A to C and A-level route—but we are also looking at graduates. We are finding, in particular, that a lot of people in this middle group are not really clear where to turn or what pathways there are. Some are feeling that the schools are pressing them to stay on and they do not have key knowledge of the other routes that are there.

We are also getting quite a lot of people who are quite proactive in going to college for some of their time, and it is not quite what they wanted. Some are transitioning quite a lot between agency and other jobs in quite a rapid fashion. With this information we are able to

provide insights into the lived experience of what the transitions are like, and I think some of this information can provide context for some of the quantitative information—so I am very much of the view that having the quantitative and the qualitative alongside each other is particularly powerful.

We are finding as well that within this group there are quite a lot of people who are quite proactive and are trying to give out their CVs, get employment and try different things, but they can be stifled by the lack of success they sometimes have and the lack of response they receive. Those are some of the issues we are finding.

We are finding as well that people are engaged in multiple activities—being at college, volunteering and working in one or more jobs that may be fairly non-standard and quite precarious. The other thing that is coming out is that, compared with the graduates we are interviewing, the younger people in this middle group do not really have clear access to guidance and information and are much more reliant on their peers, parents and friends to direct them. That is a bit of a preview of some of the things we are finding.

Professor Sandra McNally: Our priority is to use the administrative data from the national pupil database, which is linked to the further education individual learner records and HESA, and to link that, in the future, to earnings. It is great that we are going to have this link to earnings and employment because we are going to be able to look at these things in a far more contemporary setting: how do people doing vocational education now perform in the labour market; what are their chances of getting a job? These data are not yet available to us, but we are hoping they will be available to us in June next year. My biggest concern is that the data are made available to us soon because we have three years of funding and I want to make sure that our research is as good as it possibly can be in the time we have funding for. The other thing I would say is that we are in a rather privileged position of being funded by BIS as a research centre, meaning that we are going to be able to access this linked data set without any constraints.

Other people have to be funded by BIS or DfE in order to be able to access linked earnings/education data. You can easily fix this, and it is very important that you do, because you want the social science community to be using these data sets, even if they do not have funding for a particular BIS or DfE project. That is one of the most important things I wanted to bring to your attention. The ILR data set is extremely complex to use, much more complex than the schools-level data. Part of our contribution is to try to code those data better and

help make them more accessible for ourselves and other people in the future, which we are doing.

The complexity is not really the fault of the people who are collecting the data; it is partly to do with the system itself. If you are on the academic route in education, you have a very clear progression route of GCSE to A-level to university, and that is very easily understandable. However, only 40% of people do A-levels. The other 60% do something else, and it is harder to classify what those people do. They can be doing all sorts of things, and the coding is extremely complex because the system is complex. So we have spent many months trying to organise the data so that there is one learner per record in the data and to classify how you would describe what those people are doing. That is an issue which is broader than the data and goes into what the system is like itself.

Q142 Baroness Stedman-Scott: Do you find the data available in a form which make them useful for practitioners, researchers, young people and their support networks?

Professor Sandra McNally: The data have come to us in a form that we have found very difficult, but we are getting to grips with things now. It is part of our skill set and role to be able to use those data sets and try to put them in a more usable form. We have very good relationships with the analysts and have been working quite well together on those things.

Dr Stefan Speckesser: May I add a point on this? On the ground it is a fragmented scene when it comes to who generates the data, in a way. Obviously, the schools have destination information that they give to local authorities, who collect whether anybody has no destination after the summer holidays, because they have a duty to follow up on those youngsters without a destination. The local authorities collect a lot of information on young people's early post-key stage 4 activities—whether they go on to further education college, or whether they go on to an apprenticeship or to sixth form. This data set is key to understanding the participation and is more important than the school-level data to understand what people do, because, basically, the precarious biographies can be found only in such data. It is important that the development of data that is out there in various parts of the Government—in the Department of Education, in the Department for Business, Innovation and Skills, in HRMC on earnings—continues so that the data quality improves and we have better information about what young people do, so that we can see their initial situation in the labour market, from where we follow what they then do in terms of education and work. So there is a lot that can be improved in the data infrastructure.

Professor Anne Green: There is also an issue that at local level, outside the research community and universities and consultancy organisations, there is dependence on the skill sets and experience of the people at local level for the use they may make of some of these data sources. Some of the cohort studies, and the Labour Force Survey, can be readily accessed via the Data Archive, but, as has been mentioned, some of the linked data are much more of an issue, as Professor McNally has just outlined. So there are issues about skill sets to be able to make maximum use of the data, and time and availability once you get out to local-area level.

Baroness Morris of Yardley: I am still pondering over answers to previous questions. I want to try and link it to what you have said, as to why BIS will only let BIS-funded researchers use its database. Is it the complexity we have been talking about that means it is nervous about people's ability to use it? I am trying to think of a justification for that.

Professor Sandra McNally: BIS will allow people to apply to the individual learner record database and link that to the national pupil database. That is not an issue. The issue is the earnings part of the ILR/HMRC data. The linked data set on earnings can be made available to researchers only if they are funded by BIS or DfE.

Baroness Morris of Yardley: What is the reason for that?

Professor Sandra McNally: A legal reason, apparently; it does not have to do with BIS not wanting to.

Baroness Morris of Yardley: Do you think people at BIS want to solve that problem?

Professor Sandra McNally: I think they do, yes. I have spoken to civil servants in both DfE and BIS and I think they would like that to be resolved.

Q143 Earl of Kinnoull: Perhaps I can probe that a bit more. We were hearing last week—I have the quotes in my book—from people from the three different departments saying there are data protection legislation problems, but often people clearly were not qualified to talk about data protection issues. Also, there was evidence of at least one department being simply unwilling to share the data it had. I wonder whether you could comment a bit further, and I would encourage you to be very open here. It is really a question for all three, although I am looking at Professor McNally. Do you feel that there is this resistance? Could changes be made to data-protection legislation? Are you in a position, in your admirable unit of the LSE, where you cannot actually share with other people a lot of the information that you find out from the data because they are not within the envelope that you are in?

Professor Sandra McNally: To deal with the last point, our plan is to make our coding available and to make the data easier to use for people who have permission from BIS or DfE to use the data, but we will never own the data ourselves, so we will not be the person who says yes or no to other researchers who want to use the data.

My experience of BIS and DfE is that they are good about making the data available, and they do say they want to make the linked data available. I hope they can do so as soon as they say they will, in June, for the earnings data linked to the education data. That is my primary concern. I do not doubt their good intentions in that respect at all.

My experience of data with other departments is limited. I cannot really say what DWP, for example, or other departments are like to interact with on data. I think it would be a great development if administrative data across government could be linked, as indeed it is in other countries, in Scandinavia for example, which would really allow us to find out more pressing questions quickly. There seems to me no good reason why that should not be done. I appreciate you have to protect people's privacy and have data protection there, and you have to make sure that the institutions and people are using the data in an appropriate way, all the right security arrangements are in place and they are using it for the right purposes. However, once you have those things set up, it seems to me really important that researchers are allowed to use the data so that we can answer the questions that you are interested in quickly and well, and do cutting-edge research. We do not necessarily need to be funded explicitly by, say, a BIS grant or a DfE grant to do that work. A lot of people would willingly do that in universities. It is a question of making the data available and putting in place the proper arrangements. It is wonderful to see the national pupil database being linked to earnings, but I hope that the conversation does not stop there and that we can be more ambitious about what we can do.

Dr Stefan Speckesser: Perhaps I could add some smaller points. Obviously, the legal basis of the Data Protection Act, with which we all have to comply, sets out clear regulatory requirements on how to handle such data when undertaking research. So I believe the processes are probably all in place to handle such data in the research community, as it is done in other countries, for example in Scandinavia. To find more answers out of available data in public administration, the investment should be in a clear legal framework for data from different government departments to be linked, to answer questions that we currently cannot answer.

There are many other questions, apart from earnings and employment outcomes, which we could, in principle, look into—well-being, health, and crime—as an outcome. The international evidence base shows the crime-preventing mechanism of education, and this is an important element of the research that we cannot currently represent.

I believe the regulatory requirements to bring such data into a research community are very tricky, because it is obviously very personal, sensitive data, if we talk about people in prisons participating in learning. But, ultimately, a larger picture of what can be done with the data emerges, and I believe that if we progress working towards this, it can give us a better evidence base.

Earl of Kinnoull: Can I clarify one particular thing and put on the record that the ultimate genesis of the Data Protection Act 1998 was an EU directive? The same directive would apply in Scandinavia, yet somehow they have managed to transpose that directive into their domestic law in such a way that academics can work in safety, but we have managed to transpose it here in a way where it is more difficult. Is that what you are saying?

Dr Stefan Speckesser: To be honest, I am not sure that it is a matter of the Data Protection Act mechanism, as such. It may be a matter of the different government departments' remits on how to share data internally to facilitate the creation of data sets rather than whether or not you can legally access it under this mechanism. I am not a lawyer, I am an economist, but I suppose the regulatory framework in Scandinavia is equally compliant with the European directive as it is here. There may be options for facilitating this under existing regulatory frameworks. It is more of an institutional issue and an issue of the resource that would have to be considered to set up such resources, as in Scandinavia, which probably would materially change the situation.

Professor Anne Green: Just briefly, we can think of three themes here. There is the legality issue, the staffing issue and the understanding issue. We have made progress on the understanding issue. It is the staffing and regulatory issues which seem to slow things down more as opposed to the understanding issue. Obviously, for the sort of work you are interested in, it is an area and age group where you do have several different government departments all having a role, from DWP, BIS, DfE and all the rest. That makes it all the more complicated when you come to interdepartmental issues. That is a more general point, I think.

Q144 Baroness Morris of Yardley: Given our report, I want to be clear we have understood the data on the vocational route participation. You answered a lot of this question in some of your comments earlier on, so I wonder if I could ask for clarification. A young person in that age group going through the system leaves school and, let us say, might start an apprenticeship and might stay on the apprenticeship. What is the chance of being able to follow their route from school to employment? What is the chance of being able to monitor and have data on the route to employment of a youngster who, perhaps, does not go via an apprentice route and then drops off, so their route is a bit more varied? You did say it was getting a lot better, which was heartening.

Professor Sandra McNally: We will be able to do that when we get the HMRC data linked. We cannot do it now. We can only follow them up to age 20 or so, into higher education. We have looked at the NCCIS, which Stefan has been using, which local authorities have to keep. In practice, we have found we cannot follow people after about 18 because they get lost. It is a reason why we need the HMRC data and, preferably, more information than that. For example, it would be useful to know occupation. That would require ONS linkage to the data. That is not standard. Even with the data we are going to have we will not have that information in the data, and it is another reason why we need more things to be linked together.

Baroness Morris of Yardley: You mentioned local authority data and we have been told about the responsibility of local authorities to collect data when young people leave at 16. You are not referring to that, are you?

Professor Sandra McNally: It is that.

Dr Stefan Speckesser: Maybe I can give you an example of how we look into people. We are currently working on a project looking into one cohort of 16 year-olds leaving after GCSEs, which is basically the end of secondary education. Formerly, it used to be the end of compulsory education, but that does not apply any more. We are currently looking at the leavers' cohort from secondary education at census level, so everybody in the year 2011. We have been able to follow these people for about four or five years now, practically. The data we have, to understand what they are doing, is the information collected by the local authorities about their initial September status. Then we have information on whether they participate in key stage 5 or further education. Then we have information eventually on whether they go into higher education or continue in the further education sector. We can

already look into achievement and progression in the qualifications they acquire. The problem we have is that we cannot clearly say what the labour market trajectory coming out of that is, because, for that, we obviously would have to know their employment status further down the line—hence we need the merged data.

From the merged education data already available we can see how people participate in compulsory and post-16 education generally, what sort of learning they participate in, their achievement, whether they progress within the system and what the pathways are from, say, initial further education into apprenticeships or other valuable programmes.

If we focus now on vocational trajectories only, this is already a quite complex world in itself, because if you are navigated through that system by a very good college to an appropriate further education format, if your GCSEs are too weak, they would start you on level 1 in some field. Equally, the regulatory environment changes. We now want everybody to be better in English and mathematics so, if the GCSE results are too low, they will have to re-cover some skills in that direction, and that changes the participation pattern.

By and large, there are different trajectories. People may start on low levels and they may or may not progress to a level 2 qualification, which is where a lot of people drop off. The colleges aim to try to retain people to progress on to level 3, if their level 2 is suggesting they should, because there is a higher benefit, but at the same time we see more complexity; people may change subjects and may side-track at the same time, or they may start at level 2, it is going to be difficult and they go back to level 1, which could be a useful decision to progress to level 2, ultimately.

So if you look into 600,000 people leaving key stage 4 in 2011, which we have done, you see that there are very, very different patterns emerging; some are pre-vocational progression to vocational; some re-cover GCSEs; some go to level 3 directly because their GCSEs are good enough, and they may or may not go on to higher education, the kind of non-standard, non A-level way, and the FE colleges have a clear role in this. All this complexity has never really been described very clearly in terms of the economic meaning of all that. Why do they do it? Some of what we do in our world is to provide clearer descriptions of what actually happens to these people.

The missing bit is the non-education participation. For the initial September after key stage 4, most of the cohort are seen in local authority data. The problem is that local authorities should always follow up on people until age 19, so if they leave further education college

there should be a destination and a follow-up from local authorities on whether these young people actually made a successful transition, whatever they are doing if they are not working. Currently, we do not know this because the local authority collection of data is not good enough, and our data on employment are not there yet. Even if they were there, young people's earnings are often very low, so they may work and you do not find them in HMRC data. If they do apprenticeships, the minimum wages are too low to even go to the threshold.

These are real issues in describing what is going to happen, but I believe, after five, six or seven years of work, I have learnt a lot more now in what I can say about young people in their different trajectories within vocational education at different levels of skills: better GCSEs, a higher level of learning and longer participation in programmes; lower-level skills, a more fragmented pattern of participation. We have to understand what drives this participation and how we can facilitate it leading to something useful.

Baroness Morris of Yardley: That was really helpful. Thank you.

Professor Anne Green: Can I add something about keeping in touch with young people? We can learn something from some of the cohort and panel surveys as well, which are complementary to this administrative data, but we do face, with this group, the fact that this tends to be a group that is geographically mobile and difficult to keep in touch with. We know that surveys, such as the Labour Force Survey, have problems in getting responses from, in particular, young men in large cities. So perhaps some of the groups you are probably most interested in are the most difficult to capture in any of these data sources.

Dr Stefan Speckesser: Which is where education can make the biggest difference, I would add.

Professor Sandra McNally: To add to this conversation about people, particularly people at the lower end who drop in and out of education, it does suggest that it is important to link the administrative side to benefits data—DWP, presumably, has that—and, also, crime data. We should certainly find out if education prevents people from going into illegal activities. But we need the data to be able to look at that.

To add to the issue of particular gaps in the data, you probably have heard before that UCAS is one of the big missing links in the administrative data. This mainly applies to people who would be applying to higher education, but it is very annoying from a research point of view that we do not have that link, and I do not see why UCAS will not provide the data.

The other problem is that a lot of public funding for further education goes through private providers. There are fewer good data available on private providers than public providers. They are getting public money, so they should have to provide more information about themselves and it should be more accessible to us. We are finding that very difficult.

The other aspect is teachers. We do not have good linked teacher information anywhere in the system—not in schools or in further education colleges. The major public expenditure in education is on teachers and we should be able to link that for research purposes, but we cannot.

Baroness Morris of Yardley: Do you mean information about teachers?

Professor Sandra McNally: What teacher is linked to what student, or what class. I can see why teachers may not want that to be used, for example, for assessing their value added and paying them—and I do not agree with that, either—but for research and analysis we should be able to link the information together.

Baroness Morris of Yardley: That is a different ball game. I am thinking about secondary level, where any child in any one week could come across 10 academic teachers as well as two or three pastoral teachers, and wondering what useful conclusions you might be able to draw from that.

Professor Sandra McNally: If you were able to link the teachers in core subjects, such as maths, English and science, which everyone has to do, you could maybe link teacher characteristics, such as qualifications and experience, to pupil outcomes. How important is it for teachers to be qualified in maths or to have a 2:1? How important is it for people to be well paid? How does that link to student performance? Those are the sorts of questions. If you see the same teacher with different students, how important is that individual? That sort of thing.

Baroness Morris of Yardley: I agree with that. You could do that by comparing pupil achievement data, for different teachers, could you not?

Professor Sandra McNally: You could do that. If we had some identifier code—we would not need to know who the person was, as obviously that is anonymous—and if you could link the teacher to the class, even just for core subjects, that would really help a deeper understanding about that relationship.

Q145 Baroness Stedman-Scott: You have been very helpful to us. I would like to ask you two things. First of all, when you talk about private providers being very difficult in terms of

not keeping data, et cetera, the driver there could be that they only keep what is in their contract, and what they are paid to do. If it is not there they will not do that. The other point is that once you have achieved an outcome for somebody—you have found them a job—then that is it. It is enlightened providers, and there are some around, who will continue to track those people afterwards and have independent verification that the outcomes are being sustained. The best way to get those people to respond is to offer financial incentives, such as vouchers. It is costly but it gets powerful information, which is interesting.

Professor Sandra McNally: I take your points, but I also think if you could follow people in administrative data you would not need to worry so much about whether institutions can follow their students or not, and could be saving money in some ways in not pressing providers to actually track students.

On the private provider issue, we should be able to get things such as how much funding goes to private providers exactly. What is published by the SFA refers only to contracts above a threshold, for example. We want the whole thing. I do not know whether or not that is available; that is a question I have asked BIS. You would think things like that would be easy for us to get, but they are not that easy for us to get. There are lots of other things we might want to know about private provision, given how much public funding goes through them. I do think that they should be obliged to submit a return to BIS or to the SFA for getting public money, in the way that schools and further education colleges have to.

Dr Stefan Speckesser: The destination information coming through administrative data has improved dramatically over the last couple of years. When I started looking into these, assuming there would be something informative, it was very sparsely populated, but obviously the colleges and private providers are improving in supplying us with such information, so it does deliver a lot of useful information now.

I believe one of the core problems is that six months after leaving education you cannot say very much about whether a trajectory has settled, so, to some extent, some more long-term tracking has to happen. One way or another, you will not be able to facilitate that by approaching providers or colleges. To some extent, we have to have independent research on looking into records to see whether this is going towards something sustainable.

The other thing is what unconditional destination statistics tell you in areas where there is a bad local labour market. That may have nothing to do with people's characteristics or education success; it can be a matter of what is the surrounding area effect.

In a way, I believe a lot has happened for the better, to be honest, and we have to do more to contextualise that and bring it into some meaningful narrative rather than looking into a cohort leaving, and then having success or not. I believe that the success is materialising in the adult employment trajectory: that is what we have to look into, rather than destination statistics.

Professor Anne Green: One of the challenges referred to is that these transitions are taking longer, and young people are taking longer to settle. As you say, a snapshot of a few months might be somebody trying it out and thinking, “This isn’t for me”, and then going on to something else. Transitions are not always unidirectional; it may appear that they are taking a step back to go forward.

We also have to keep in mind that there are some things we cannot really get a handle on even through linking to admin data. For example, we cannot assume that people who are not in employment or education are going to be on JSA or a benefit, because some people’s parents will feel that they want to support them in that period and will help them out. So they are hidden from the statistics in any case.

This point about needing tracking for longer is a very important one, given the contemporary transitions that we are finding, as is linking to the deprivation data at the neighbourhood scale and data at a more local labour market area level, to look at the issues of tightness or slackness of the labour market and how different groups of people fare according to local labour market conditions.

Q146 Lord Farmer: Coming to demographics, there are likely to be differences in patterns of participation and progression for young people from different backgrounds who do not follow the academic route. Is data analysis currently being undertaken that offers a breakdown by different demographic factors? What does this tell us about differences in patterns of participation and progression?

Professor Sandra McNally: There has been a fair amount of work on this, and there is current work being undertaken on this which will develop. Before I pass over to Stefan, who has done most of the work, my big take from it is that family background and prior attainment are the most important things. That is it.

Dr Stefan Speckesser: This is, basically, the finding of a recent research project for BIS, which we had started under the assumption that we could provide some long-term outcomes of particular pathways. We found that these data were insufficient. We then did a systematic

literature review using all the published econometric research on the topic. There is very clear evidence about certain demographic factors influencing participation as well as attainment. They are the usual suspects. Good GCSEs help you to progress; there are gender and ethnic group barriers that are obviously visible; there are local area effects; and then there is a whole area of parental background and traditions in the family. It is not only income or social status; you may also find some vocation orientation in the family's trajectory. The truth is that the administrative data will not be extremely helpful in shedding more light on this.

In the administrative data, we have a clear indication of a household's income, because there is free school eligibility information, and that shows whether people come from a poor background. You also have information about the local areas and deprivation that you can link via postcodes, or other mapping. You can aggregate these data to school level, and that can show the percentage of free school eligibility in a school. That gives information about the sort of area environment in which education participation happens. In the ideal world of a researcher, you would like to have more tangible family background characteristics, one of which I believe you cannot really get out of data unless you look into cohort studies where there is very well-documented parental information. Are the families actually positive about vocational education as a trajectory, or is it the mainstream, everybody wants to go academic anyway? There may be different lines of segmentation rather than income or local deprivation that can influence trajectories if families have traditions or they have certain businesses. They may have a taste for vocational education or they initiate vocational education, rather than seeing it as a second-best solution as opposed to high academic achievement. Unfortunately, we do not have much information about such data. I guess there is a lot of qualitative research about this. We probably can use some of the research data from Essex, or wherever it is archived. For example, the occupational codes of the parents could give you some information. I have not seen very much on this yet and we probably should complement administrative data by such a component, if possible.

Professor Anne Green: The cohort surveys and the panel data do complement the administrative data to get those sorts of insights, but when you have a cohort or a panel that you want to break down by gender and ethnic group, and all the rest of it, the sample size numbers are often relatively small, which can be an issue. But the qualitative work that we

are doing does back up what you were saying about parental background. The openings through parents, friends and siblings are very important for this particular age group.

Lord Farmer: It sounds to me as though that is an important area in which data should be collected, yet you are saying that it is not. So what can be done?

Professor Sandra McNally: There are some excellent cohort studies: the Millennium Cohort study, for example. These cohort studies, and the Labour Force Survey, are all complementary to the administrative data sets, and the question is, how do we get the most out of it all? In an ideal world, you would want to link the cohort data to the administrative data. The difficulty is that you need explicit opt-in consent by participants in order to do that, and that is a pain. We would much prefer opt-out consent rather than opt-in consent—it would make our lives so much easier—but that is apparently in the Data Protection Act. If we can do that, it is important.

One thing I would like to highlight is that sometimes BIS and DfE commission researchers to do surveys and the raw data are not necessarily deposited with the government department. An improvement would be to make researchers who collect data on behalf of any part of government deposit their data in the data archive as a condition of getting the grant. That is probably an easy way to make data more available to others to do subsequent research.

Baroness Morris of Yardley: Can I go back to what you said about your findings, which is what we expected, that parental background and family background, home circumstances, made a difference? When you say that, do you mean their socioeconomic background? We believe that aspiration makes a difference as well. You can be poor and still aspire for your children to do well. Do you have any way of talking about aspiration separately from family income? I worry sometimes that we associate low family income with low aspiration. Although it might be truer now than it was two generations ago, it is almost labelling people in a way that might not be helpful to change that situation, even if it was the case. Do you ever make that difference? I realise it is difficult, quantitatively, but I wonder if it is something that might be seen in a different form of research?

Professor Sandra McNally: I think you can in some surveys which have information about aspirations as well as background. The reason we focus a lot on background is that you can observe it in all sorts of data sets, as well as international data sets. It is objective; we can see if somebody is on a low income, or poor, or on free school meals, and we know exactly what that means. The troubling thing about our findings in the UK is that it seems that family

background is more strongly linked to progress in education and labour market outcomes than it is in other countries. That is a big issue for policy. At the same time, it would not be saying that for people from low-income backgrounds it is a sentence and there is nothing that can be done. Clearly, there is some heterogeneity there, and some people from poor backgrounds have high aspirations. That is something you can see in research as well, but I am not quite sure what you do about it. I would not come to the conclusion that all you have to do is raise the aspirations of poor people and that is your problem solved.

Q147 Baroness Blood: Thank you for the information you have given us this morning. Could I change tack here and talk about the role of the Government? Could the Government do more to ensure that information about this middle group of learners is shared? Or do you think you are where you are because of what the Government have done in moving the situation on?

Professor Sandra McNally: My priority as a researcher is to make sure that the raw data are shared with us so that we can do some analysis, produce findings, and disseminate those findings to a broader public who can use it, such as policymakers or practitioners. There is another issue of parents and local authorities and teachers, and how you make data more accessible to them. It seems to work better in some parts of the system than others. For example, it is very easy to get information about schools; you can get Ofsted reports and you can get data on schools. I do not think that would be as true for FE colleges and private providers, for example. So there is probably more that can be done there, but the starting position is to get the better raw data out to researchers—but the other issue is important.

Dr Stefan Speckesser: Could I add one point? As I said earlier, the local authorities' duty to follow up on young people should result in a good database on what activities are undertaken. We all know that these data are not collected very consistently, and there has been movement to improve it, but this is a moving target. We have to continue working towards better evidence on what young people do. I believe that that has to be interconnected better with the education institutions on the ground, because information is not very often shared between FE colleges and schools. Schools know a lot about parental background and how pupils have performed during compulsory schooling, and that information is not handed to FE colleges either, for example, so it would be easier to access the data not only for researchers but also for the people who run education programmes on

the ground, to have better information on the people who are participating in their programmes.

Q148 Baroness Stedman-Scott: Do you think that the collection of data could or should be linked in a productive way with incentives on performance measures for schools and colleges?

Professor Sandra McNally: I do not have a strong view on this question, except to say that however the data are collected should not give providers the incentive or opportunity to cheat or misreport findings. I am a bit suspicious of linking things too much to incentives.

Baroness Stedman-Scott: That is a fair point.

Professor Anne Green: Our experience when we have done surveys of young people themselves is that incentives do work; entry to a prize draw can be helpful. That is at an individual level as opposed to a school or college level. These things can work to some extent, but are probably not the whole answer.

Q149 Baroness Howells of St Davids: Do you have one key suggestion for change that this Committee could recommend to improve the quality, quantity and application of data in policy related to school leavers?

Professor Sandra McNally: My one recommendation would be to make administrative linked data as efficiently done across departments as possible to make it available to researchers who have the right institutional requirements and have gone through the security checks, and to make access to that data available to those people regardless of whether they have funding from BIS or DfE to do a particular project.

Dr Stefan Speckesser: I can only say I agree.

Professor Sandra McNally: We talked before.

Dr Stefan Speckesser: She is the director, after all.

Professor Anne Green: It sounds sensible to me as well—and she did not tell me that beforehand. That would obviously be a common-sense way forward. I cannot disagree with that.

The Chairman: Thank you very much for coming today and sharing your expertise with us. We have found it very useful.