



Public Administration Committee

Oral evidence: Statistics and Open Data, HC 564

Monday 11 November 2013

Ordered by the House of Commons to be published on Monday 11 November 2013

Members present: Mr Bernard Jenkin (Chair); Alun Cairns; Paul Flynn, Kelvin Hopkins; Greg Mulholland; Lindsay Roy; Mr Andrew Turner

Questions 131-248

Witnesses: **Michael Fallon MP**, Minister, BIS and **Mr Nick Hurd MP**, Minister for Civil Society, Cabinet Office, gave evidence

Q131 Chair: May I welcome our two witnesses to this third session of PASC's inquiry into statistics and open data? Could I ask each of you to identify yourselves for the record?

Mr Hurd: Nick Hurd, Minister for Civil Society.

Michael Fallon: Michael Fallon, Minister of State for Business and Enterprise.

Chair: Thank you very much indeed for coming.

Q132 Greg Mulholland: Good afternoon, Michael; good afternoon, Nick. I will start with a very broad question, to set the scene, looking at the overall aims and objectives of the Government's open data strategy. There seem to be three overarching objectives that the Government is trying to achieve through the open data strategy: first, to increase accountability; secondly, to improve efficiency of public service; and, thirdly, to promote economic growth. All members of the Committee would entirely support all three objectives, but is it realistic to be pursuing all three? Where there are conflicts between them, as there may be, which one takes priority?

Mr Hurd: Thank you very much for inviting us. It is entirely realistic to pursue all three. I would argue that it is extremely necessary, because of the value and the benefits that we think we can unlock in changing the culture of Government from one where effectively information and data was hoarded to one of openness by default. That is what we are trying to achieve, because as the Minister for the Cabinet Office has said, we think that open data could be one of the new raw materials of the 21st century, with huge potential, as you say, to deliver all those three benefits. We have got to be ambitious. We are world-leading in this area, and I do not think we are in any mood to give up on ambition at this stage.

Q133 Greg Mulholland: Do you not feel more accountability, for example, could actually compromise the goal of achieving economic growth? I accept that this is the culture that the Government set out to challenge, but the culture of the past, particularly on the economic figures for Governments over many years, has very much been to seek to put in the public domain economic data that is helpful to them rather than necessarily the full picture, even if it contradicts certain policy decisions they may be making. Do you not feel that could lead what you are doing to come into conflict with what Departments are pursuing? I would be interested to get Michael's view.

Mr Hurd: Government sits on a lot of data that is potentially very valuable to lots of different people, and therefore the starting principle is that we should get the data out in a form that is accessible and available to everyone, and put it out in a way that is machine-readable, so they are free to use it or reuse in any way they like. That is what open data is. The starting point is, "Let's not sit on it; let's get it out there, and let's not necessarily hold that process up". We hear lots of excuses for why the data should not be out there. Let us get it out, because this could be enormously valuable for the country. We are at the start of that process. There are all kinds of complications, all kinds of difficulties and all kinds of tensions we have to wrestle with, but the starting principle of open-by-default is absolutely the right one to start with.

Q134 Greg Mulholland: And you have a genuine commitment to getting that information out, even if it is not helpful to certain Departments and to certain policies being pursued by the Government.

Mr Hurd: The starting point set by the Prime Minister is that we want to be open and transparent, not necessarily in a process of starting off by making judgments about what is helpful and what is unhelpful. This is public data; it should be out in a form that is accessible and usable. That is the starting point of principle, which I think is an important one to stick to.

Q135 Greg Mulholland: Michael, from a departmental point of view, there are tough decisions being made by BIS. What if some of the information was not helpful to policies that your Department was pursuing?

Michael Fallon: I would like to think of some examples of the kind of data that we would even be able to withhold. I have one advantage over Nick here. I was a Minister 20 years ago and there has been a cultural change in Whitehall. I think you were right that the Government did hold on to data as a principle, and unnecessarily. What has changed in the intervening years is that this is a huge emerging market, and there is really no reason why Government should have a monopoly over the use of data or over its potential applications. There has already been a cultural change, and I cannot think of examples where we would be circumscribing ourselves by holding back data.

Q136 Greg Mulholland: So you and BIS would be perfectly happy with full publication of data, even if it showed that policies being pursued by BIS currently were not succeeding?

Michael Fallon: We have an independent statistics office in this country; we cannot trespass on their independence. We are committed to the data that they publish. There are different classifications of that data, as you know, and there are definitions of national statistics and so on. You would need to give us some examples if there was somewhere, in your mind, that a series might be unhelpful and that making policy might be made more difficult because of our obligation to get a series out there.

Greg Mulholland: I will not dwell on that in the Committee now, and I will pass back to the Chair, but there probably are some examples. I might take those up with you directly, if I may, but thank you for those answers.

Q137 Lindsay Roy: Michael, can you just clarify what the essence is of this cultural change?

Michael Fallon: When I was first in government more than 20 years ago, Government regarded data as its property. There is an understanding now that that data belongs to all of us—it belongs to the public that we serve and it belongs to the taxpayers who have paid for its collection. That has been a big cultural change.

Mr Hurd: If I may add to that, Chair, there is also a change in terms of increased perception of the value of this data. 10 or 15 years ago I am not sure we understood the value of it, because there was not necessarily an information marketplace; there is now. Government is gradually waking up to the fact that we are sitting on something valuable and maybe we should share it.

Q138 Chair: This idea that pushing this data into the public domain and giving the public access to it is good for economic growth and good for economic opportunity is a very strong theme of the Government's thinking on this subject, Mr Hurd.

Mr Hurd: Yes.

Chair: Do we accept that business will be hesitant to invest in the infrastructure to make use of this data if the Government's protocols and releases are not reasonably coherent and predictable?

Mr Hurd: Businesses always want a clear framework within which to operate. That is a given. One of our good achievements is to have set up an architecture of challenge and advocacy whereby those who want to use the data—those who are advocates for this—have very good access to Government to tell us where things are not quite right and can be improved. Certainly we are not detecting any lack of

business appetite or interest in this area, at the start-up and SME end, at the more mature business-to-business end, or among some of the larger companies in the country.

Chair: So there is a need for a clear framework.

Mr Hurd: There always is when you are dealing with this.

Q139 Chair: Who is in charge of this framework in the Government?

Mr Hurd: The Cabinet Office leads on transparency and open data.

Q140 Chair: So why is Mr Fallon sitting here? There was some confusion about whether Mr Fallon or Mr Willetts should be the Minister from BIS. Why do we need two Ministers here?

Michael Fallon: I am here by invitation, Chairman.

Q141 Chair: Well, the Government was invited to submit a witness and we have got two witnesses, which does not suggest that leadership is focused in one Department. It is nothing personal.

Michael Fallon: Let me be clear about this. This is an area where the Cabinet Office leads, both on policy for open data and on transparency. At BIS, we play a supporting role, in considering our capability as a country to make the best use of data, including the skills and infrastructure that are needed. We play a supporting role.

Mr Hurd: If I may add to that, Chairman, I know the Committee has done its homework and has a sense of just how much work is involved in this, not just in gripping the culture in central Government but also in the interface with the private sector and trying to encourage innovation and increase the capability of the country. BIS plays an enormously valuable and important role in supporting us in those two critical areas: innovation and interface with the private sector; and capability.

Q142 Chair: The Government appointed an open data tsar to report on this matter, Mr Stephan Shakespeare, and he wrote in his memorandum to this Committee: “I propose that we should have a plan that clearly defines how, when and what data is released; how it will be maintained; and how it can be used, and someone to drive this forward in a publicly auditable way”. He goes on to say: “Committing oneself to transparency is not enough. Merely throwing open some cupboards, as it were, and saying to entrepreneurs, ‘Here you are. Make it happen’ is unlikely to lead to anything

but a painfully slow development of value from the data”. What is your reaction to that?

Mr Hurd: We were in broad agreement with almost everything in Stephan Shakespeare’s report and have accepted most of the recommendations and acted on them very quickly, not least in simplifying the governance arrangements. My main point in reply, Chair, comes back to what I have already said. Part of what we have achieved is to set up institutions, boards and forums that really do help to keep Government’s feet to the fire. We absolutely need that, because although lots of independent surveys and authoritative voices regularly cite the UK as being a world leader in this area, we do need people like Mr Shakespeare and reviews like his to keep pushing us, keep our feet to the fire and keep prodding us to improve. As I said, we are very much at the start of a journey here. He did also, I think, say in his oral evidence that he did not see any signs of Government Departments not co-operating; his main frustration was about pace, which I quite understand and fully encourage him to keep on about.

Q143 Chair: What he says is, “There are many committees, boards, overseers and champions of data; but no easily understood, easily accessed, influential mechanism for making things happen. There should be a single body with a single public interface for driving increased access to public service information.” Is that a recommendation you have accepted?

Mr Hurd: What we accepted was the recommendation to effectively merge the two boards, and the Transparency Board is a very effective forum for driving this forward, which is chaired by the Minister for the Cabinet Office with great rigour indeed.

Q144 Chair: The Transparency Board under you, Minister, in the Cabinet Office, is where the single strategic leadership rests.

Mr Hurd: That is where I see it, and also in the transparency boards in each Department. We have created an apparatus here of accountability, pressure and advocacy, which may not be perfect but is effective, not least in making sure that the voice of users of data is heard round that table, through the Open Data User Group.

Q145 Chair: So that is the focus for accountability for the Government’s open data strategy.

Mr Hurd: That is the body that I see driving the strategy. On the accountability that you are talking about, you can see in the quarterly written ministerial

statements the very open process of assessing the progress of Government across the Departments and a regular update on the implementation of the strategy.

Q146 Chair: What about what the Open Data User Group describes as, “A disparate legislative framework with responsibility spread across multiple bodies” leading to a pace of delivery that is “relatively slow”? What do you say to that?

Mr Hurd: What I would say is that our process is enormously aided by having lots of advocates and champions of open data who are constantly pushing us to go further, faster and higher. They are entirely right. It is part of the reason why this country is the world leader in this process. Certainly as far as the Cabinet Office is concerned, it is leading the effort to drive this culture change, and we should not underestimate the difficulties and challenges underpinning that. We need people like the Open Data User Group and the Shakespeare Review to keep telling us where we can make more progress. But let us not lose sight of the fact that in three and a half years we have made enormous progress and we are the world leader in this area.

Q147 Chair: This is very good, but you are encountering opposition in other parts of Government sitting not very far away from you—people who are defending the status quo and want to keep data locked up and under proprietary ownership, something that one of our witnesses, Sir Nigel Shadbolt, has complained about, saying that the Treasury is part of the problem: “I do not think they get much of the opportunity of the digitally disruptive economic abundances that can flow from data.”

Michael Fallon: I am not aware of that. I am not answering for the Treasury.

Q148 Chair: Well, you locked up the postcode data and sold it off with the Royal Mail.

Michael Fallon: If you want to turn to the postcode data, let me first of all say—and I think it helps Mr Mulholland—that we have published some 1,300 datasets as a Department. I think we are the third highest publishing Department of any in Whitehall, if that helps. If there are other datasets that can be added to that, we would be happy to do so.

If we can turn specifically to the postcode address file, the postcode address file is an integral part of the Royal Mail; it is a fundamental operating asset on which the business depends. It is the Royal Mail that collects the data and makes sure it is up to date. It is not simply a store of data in a cupboard that you can suddenly release; it is part of Royal Mail.

Q149 Chair: But many people would argue that the Royal Mail was part of the public sector and it collected that data on behalf of the Crown. It was, after all, called the Royal Mail.

Michael Fallon: The Royal Mail is still regulated and the legislation applying to the postcode address file is still there, so that is still a regulated entity within the Royal Mail. There is no difference there so far as privatisation is concerned. As part of the privatisation process, we took the opportunity, because we had a number of representations about this, to see what we could do to further open up the file to more access, and the Royal Mail has come forward with a package that I hope you will welcome, making sure the file is free for charities, that it is free to micro-businesses for their first year, and that people can look up up to 50 things on it instead of 15. They are now discussing a series of other minor reforms to the licensing process that might further improve access to that data.

Q150 Chair: The fact is that the Treasury wanted the postcode address file included in the privatisation, and those interested in the wider economic benefits of data did not want it included as part of the proprietary information of the Royal Mail. That is a conflict of interests, is it not, between the broader public benefit and an individual privatisation?

Michael Fallon: No, I do not accept that for a moment. It was nothing to do with the Treasury; I took the decision to keep the postcode address file within Royal Mail, because I was absolutely persuaded that it was an integral part of Royal Mail. It is not a huge contributor to Royal Mail's profits, but it is an integral part of its operation. The updating and collection of that data falls to Royal Mail, and it was not clear to me how taking it outside Royal Mail would have benefited anybody. It was my decision.

Chair: We are going to come to the address file data in more detail later on.

Q151 Kelvin Hopkins: Put very simply, it strikes me that the Government has sacrificed the public interest to commercial interests, and that the Royal Mail could have carried on assembling the data about addresses, but it could have been publicly accessible.

Michael Fallon: Royal Mail incurs considerable costs in collecting and maintaining this data and keeping it up to date. It is only reasonable that they should be able to recover some of those costs from the companies that use this data. The postcode address file continues to be regulated after privatisation as it was before privatisation. The change of ownership has no difference. If anything, it has made us look again to see where the Royal Mail can continue to open up access to that data, and that is what is now happening.

Q152 Kelvin Hopkins: But if it was kept in the public sector, the data could have been sold to outside interests and the benefit would have accrued to the public purse.

Michael Fallon: If the whole of Royal Mail had been kept in the public sector, do you mean?

Kelvin Hopkins: Even now, if the charges that are made for the data were to go to the public purse rather than into private pockets and private profits.

Michael Fallon: As I have said already—I am just repeating myself now—the collection and maintenance of the postcode address file is an integral part of Royal Mail’s business. It cannot be simply lopped off from it. Somebody has to go out and collect that data and keep it up to date every day of the year. It is not a task that can be performed by anybody else. It is integral to Royal Mail’s work and I saw no case whatsoever for somehow carving it out and spinning it off.

Q153 Alun Cairns: My main questions relate to Government Departments, but I want to pursue this issue about the Royal Mail. Mr Fallon, if this data was included with the sale of the Royal Mail, surely the commercial gain will have already been derived with the share price that has come out thereafter.

Michael Fallon: The business has only just been privatised, so I am certainly not going to comment on where the share price is at the moment or where it is likely to be when things settle down. The best future for Royal Mail is to have access to private capital and to have a secure long-term investor base to allow it to continue to develop its business in a very competitive market. That does not affect the way in which the postcode address file is updated or maintained; as I say, that will continue to be regulated, and those who might have wanted completely free access to this data—and I am sure there are large businesses out there that would have wanted completely free access to this data—will still be able to access it in the way that they can at the moment.

Q154 Alun Cairns: Mr Fallon, I absolutely accept and support what you are saying, but the point I am seeking to question is the suggestion that was made earlier that if that data was still held in the public sector, there would be a value to that. That value has already been derived in the sale of the Royal Mail as-is.

Michael Fallon: Yes, you could argue that the taxpayer has already had a receipt for that data, but there is a continuing benefit to the shareholders in having the postcode address file right at the centre of Royal Mail’s operations.

Q155 Alun Cairns: My question relates to something you mentioned earlier, Mr Fallon: the culture change since you were a Minister 20 years ago. Would you also accept there are some Departments that are better than others in relation to their

approach to open data? As a result, what would you say is the best Government Department, the one that responds most positively to the demands that you and Mr Hurd are making of them?

Michael Fallon: Mr Hurd is probably better to comment on that, because he sees the data from all Departments. I simply gave Mr Mulholland the score: we are the third highest in the number of datasets that we publish. I know who has come first and second, but I will perhaps give Mr Hurd the opportunity to comment on our relative performance across Government.

Mr Hurd: I simply refer you to the written ministerial statements, which do exactly that on a quarterly basis, including very helpful coloured charts full of greens and reds to indicate the progress made by each Department against each indicator. That is part of the accountability mechanism that the Chairman was probing me about. That goes out every quarter.

Q156 Alun Cairns: Does that publication have an effect? Does it work in terms of influencing the Departments?

Mr Hurd: Oh yes.

Q157 Lindsay Roy: Can you tell us what priority is given to sharing good practice?

Mr Hurd: That is a good question. Again, the process I have just described, leading up to the written ministerial statements and the dashboards that show where each Department is, is a fantastic discipline in itself in terms of the process of engagement with Departments. Part of the responsibility of the Cabinet Office team is to manage that relationship with other Departments and to be part of the process of sharing best practice. The Transparency Board is another mechanism to bring in Departments that have done particularly well or Departments that need particular challenge. That is part of what that forum is for.

Q158 Lindsay Roy: Can you tell us how the sharing of good practice is done?

Mr Hurd: As I have said, by civil servants talking to civil servants in those Departments, and through the Transparency Board. Departments come in on quite a regular basis to update the board on their progress. That is an opportunity to share best practice and put some pressure on Departments that are a bit behind the curve.

Q159 Lindsay Roy: In essence, what you are saying is there is a skills shortage in some areas of the Civil Service.

Mr Hurd: No, I did not say that.

Q160 Lindsay Roy: Is there not? We have some evidence to indicate that some civil servants do not have the skill sets required.

Mr Hurd: Not inside the Cabinet Office. I come back, Mr Roy, to the fundamental point here. This is about a culture change inside the system; therefore, it would be miraculous if this system were smooth and linear and if there were not scope to have the kinds of arguments that we had around certain cases. It would be a miracle as well if the Civil Service was completely tuned to this. There is an issue around capability that fits into the Civil Service; there is also an issue around the skill sets inside the Civil Service. But that goes for the wider country in terms of our ability to access what you might call data skills, and it is part of the challenge of how we meet this opportunity. I do not think it is anywhere near a crisis inside the Civil Service, but we are learning and we are learning fast.

Q161 Lindsay Roy: What I am trying to get at is what priority is given to professional development and involving new recruits.

Mr Hurd: That leads us into the issue of being clear about who you need to recruit—what skills you need—and what you are prepared to pay for them, which is a whole set of issues. There is an issue. I do not think it is the biggest issue that we face in terms of driving this forward. There is an issue around capability, but it is not a crisis—it is not an overwhelming one; it is a very manageable one.

Lindsay Roy: Surely it is fundamental that you have a Civil Service that can deliver.

Mr Hurd: I come back to my starting point. We are world leaders, so we are not doing badly. We are recognised by lots of independent surveys as that, and often recognised in public by the Americans as that. We are not doing badly. There has not been a capability crisis up to this point. We are number one.

Lindsay Roy: That might sound to some people a bit complacent.

Mr Hurd: No, we are not remotely complacent, because of what I said earlier to the Chairman. We know we are at the start of a journey. We are very ambitious. We are not listening to siren voices saying, “Go slower” or “Be more cautious”. We are absolutely going for it. There just is a huge amount to do. But there is no shortage of ambition; I think that is genuinely recognised, not least among some of the witnesses that you have had before you.

Lindsay Roy: Mr Fallon, do you have anything to add?

Michael Fallon: No.

Q162 Chair: What are the skills required in the Civil Service to make the best use of open data?

Mr Hurd: At one level, the thing we are most concerned about is getting the data out. That requires will more than anything else at the moment. What we are trying to construct is mechanisms whereby the Government has access to people with superb data skills and data knowledge. We have a group of very competent civil servants leading this at the Cabinet Office, but our main challenge is working with other Departments to get them over the line in terms of releasing data.

Q163 Chair: I am sorry; this “fire and forget” mentality is what many of our witnesses are very critical of. Sir Nigel Shadbolt said there was “a real challenge” within the Civil Service to find officials who understood data. Do you not accept that?

Mr Hurd: I accept that it is one of our challenges, but I come back to the point that if it was an overwhelming challenge, we would not be world number one in this area—a field where lots of countries share this ambition. I am just trying to get it into perspective, Chair; that is all.

Q164 Chair: There are lots of other countries that are going down this open data route that are taking a much more strategic approach to the development of capability within the Civil Service and the supporting of business and infrastructure beyond the state in order to create best value around this data, and yet we find one of the key players in the sector, FlyingBinary, says, “We find ... data literacy is very low across the Civil Service”. What should we do about it?

Mr Hurd: We look to continually upgrade our capacity, and we look to continually insert into our process people who know more than us. That is how we learn. We build into our process people like exactly the people you are quoting, who challenge us and show us the path. I challenge your assertion that there are other people doing this better than us, because we are consistently recognised as world number one. I am not remotely complacent about that at all, but we are not doing badly.

Q165 Lindsay Roy: Should it not be a higher priority?

Mr Hurd: Higher than number one?

Lindsay Roy: No, a higher priority in terms of skill development and training.

Mr Hurd: The commitment to open data and transparency is given the highest priority, because it comes directly from the Prime Minister. I genuinely do not think there is an issue around ambition or commitment, nor do I want to denigrate the work and effort of outstanding civil servants in this area. We know this is a

complex area. The private sector wrestles with this as well, in terms of the ability to access the skills they need. It is the same for every country. I do not think it is something that we are remotely complacent about.

Q166 Lindsay Roy: Is there not an issue about outcomes in some Departments?

Mr Hurd: You are always going to get uneven levels of commitment and performance across Government. That is just a given. But again, I come back to my point about the architecture we have set up around this—the institutions, the push and pull. That is why, with business support, we have supported the creation of the Open Data Institute as a centre of expertise and goading that helps our civil servants, helps us raise our learning curves and gives us interaction with the entrepreneurs and the people developing the apps, who know what they want and share that with us. That is why we create spaces in which our civil servants can interact with exactly those people, and that is recognised as a fantastic innovation around the world. Again, I am really not wanting to sound complacent at all, but I would just like you to recognise that we have put in place some very important building blocks, of which the ODI is one.

Q167 Kelvin Hopkins: Just following on this theme, when I was a student in the 1960s there was concern that, in Britain, Government had a deeply ingrained instinct to secrecy compared with, say, America, Scandinavia, Canada and elsewhere. I just find it difficult to believe that we still have not got a tad of that continuing in Government. We are not that keen on giving out information and data. Is that not fair?

Mr Hurd: I will make two points. Firstly, we have made a huge amount of progress in terms of the number of datasets that are out there. Mr Shakespeare said in his evidence to you that he had not come across any significant obstruction; his frustration was about pace. Yes, you will find some caution in some areas around the issues of privacy and quality, and you will find some areas where people are a little bit slower to get the vision and the value, if you like. This is a process of persuasion, but under the leadership of the Prime Minister and with the rigour of the leadership shown by the Minister for the Cabinet Office, what we are trying to establish—and I think we have succeeded—is an open-by-default culture where you publish what you can publish unless you have got very good reasons not to. You will have arguments along the way, you will have differences of opinion and you will have cases to the contrary. That is an inevitable part of the process. But in three and a half years, we have moved a very long way in changing the culture.

The next critical step will be improving this by pointing to applications and new platforms and programs set up to use the data that are valuable and that our constituents are using, so that Departments can really see the value. That is beginning to happen. That is part of our role: to make this stuff come alive in people's minds so that there is greater buy-in to the value of this data, because people are using it in ways that we in government could not possibly have imagined.

Q168 Kelvin Hopkins: Tony Blair regarded the greatest mistake he made in government as allowing through the Freedom of Information Act. I remember it going through the House. At one stage, a provision was put into the Bill to say, “You can have as much information as you like, unless we say ‘no’”. It was a heavy qualification, which I and many others on my side voted against, because we thought it would destroy the principle of freedom of information. At the moment, Departments decide what datasets they are going to release. Should there be a right to data instead?

Mr Hurd: As I said, we are in a situation where the culture is open—

Chair: That sounds like a “no”.

Mr Hurd: No. We are in a situation where the culture is open by default. One of the most critical initiatives that we have undertaken, in a direct response to the Shakespeare Review, is the setting up of the departmental datasets and the National Information Infrastructure so that people will see what we have got. Then the process is in place by which Government can be challenged—“Well, why are you not releasing this stuff? What are your arguments for not releasing it?” As far as the Cabinet Office, which is leading this strategy, is concerned, our default situation is open by default unless you have got a jolly good argument, so bring a jolly good argument. Until you have got the information out there—until you have got this inventory, which is a really radical step—it is very difficult to have that kind of transparency and that debate, and therefore power rests with the Departments. With the combination of the data strategies and the process I have set about, we are changing that.

Q169 Kelvin Hopkins: The latter part of my question was, “Should there be a right to data instead?” I imagine in Scandinavia or elsewhere they would say, “You have a legal, constitutional right to data”. That would solve a lot of problems. I have to say, I am not convinced that we have overcome the problem of this ingrained secrecy I mentioned earlier.

Mr Hurd: I think we have come a huge way, not least in the evolution of the Freedom of Information Act and the right-to-data clauses in it, but also just in the whole process we have set up through data.gov.uk and the feedback mechanisms on that. Again, I come back to the process whereby, in response to the Shakespeare Review, we are putting it all out there and making it quite clear what we are not publishing. It is open to challenge in a completely transformed world of accountability and transparency, with all these mechanisms that I have talked about such as the Open Data User Group and the Transparency Board, whereby our feet are kept to the fire and the question is always asked, “Why are you not opening up those datasets? What is your argument for not opening up those datasets?” We have changed the default, and Government has to justify not opening them up.

Q170 Chair: Mr Hurd, you have been giving very long answers to this very simple question: “Should there be a right to data?” Can you give a “yes” or “no” answer?

Mr Hurd: In effect, that is what we have created.

Chair: I just want a “yes” or “no” answer.

Mr Hurd: We do not see any need for legislation to bring that about.

Q171 Chair: So it is “no”. The Government does not believe in a statutory right to data.

Mr Hurd: We have got no plans to legislate on it.

Chair: “No plans to legislate” means “no”.

Mr Hurd: Beyond the Freedom of Information Act.

Q172 Chair: You are against a statutory right to data. That is the Government’s policy.

Mr Hurd: I am, at the moment, against any further legislation in this area other than evolution of the Freedom of Information Act and the transposition of the EU directive in this area.

Q173 Kelvin Hopkins: This is a question to Mr Fallon. BIS funding for the Administrative Data Research Centres has been widely welcomed, but is it right that no new publicly available statistics will come from this initiative, and that access to the datasets will be limited to academics? Should private and other bodies have access to it?

Michael Fallon: I do not rule that out in the future. We have just started with the Administrative Data Taskforce work and its networks. I think you could make a case for allowing access to anonymised Government-held data to inform commercial decisions and improve, for example, the effectiveness of business investment, and I would be happy to look at that a little later on. The first priority must be to demonstrate the security of the data, the robustness of the data and its usefulness, and to realise its potential benefits for our own policy-making processes. That is the first task, but I certainly do not rule out going further once we have set out down this route.

Q174 Kelvin Hopkins: I detect a note of caution, even at this stage, and some slight reluctance to be as open as perhaps we might like.

My last and probably most significant question is about the fact that we have heard evidence that the outsourcing of public services undermines open data principles because firms claim commercial confidentiality. Should the Government insist that all companies with Government contracts observe open data principles in future? Should they not just be treated like any other public sector organisation, given that they are being funded by the public purse?

Mr Hurd: The situation at the moment is that we are working with our colleagues at GDS on a piece of work to include open data clauses in IT procurement, and we have made some money available for Departments, agencies and local authorities to release data where there are short-term technical barriers—i.e. where someone is saying, “We are going to have to charge you to get those datasets out”. As I have said, we are working with our colleagues in GDS on some work to include open data clauses in IT procurement in particular.

Q175 Kelvin Hopkins: We have already seen a substantial shift of what was done in the public sector to the private sector with outsourcing, so we have lost a great area of data, which is no longer available even with FOI. I hope it reverses rather than continues, I have to say, but with further privatisation and outsourcing we could see a diminishing to a very small area where Government data is available, so much of the meaningful data that we want to know is in private hands and we cannot get our hands on it.

Mr Hurd: There is a risk, but it is a manageable risk. When Government outsources, it places requirements on contractors, and there are requirements to report, not least on outcomes, and the data flows. What we are actively working on is how we can build into that a process whereby we effectively remove barriers to open data. As I have said, we are actively looking at the IT barriers in that context and working with our colleagues at GDS on that.

Q176 Kelvin Hopkins: Is the reality not that some of these private companies and outsource organisations would just not be interested in doing the job if they had to be open and provide information and data to the public, as the public sector is required to?

Mr Hurd: It is all part of this culture change in Government that we are driving, and I think largely succeeding in driving, so that at the start of a commissioning process you are very clear about the outcomes you want and the risks that you are managing in terms of the requirements you place on your supply chains.

Kelvin Hopkins: I have finished my questions, but I would like personally to see why the privatised railways are now so incredibly expensive compared with what went before, and the massive debts of Network Rail—

Q177 Chair: I think that is a little off the case, but you have made your point.

I would like to point out to our two witnesses that there does seem to be something of a dichotomy between the rhetoric that comes out of the Cabinet Office and the very cautious hedging around all these questions that are in front of us today. Is that why you are here, Mr Fallon—to stop the Cabinet Office from giving away too much?

Michael Fallon: That is an extraordinary suggestion. I have already said that the Department of Business has a very good record in releasing its own data. I also have ministerial responsibility for trading funds that have very formidable sets of data between them. We may be coming on to discuss those. That is an area we work very closely with the Government on.

Q178 Chair: Coming back to a question that Mr Hopkins asked, why should private and third sector bodies not have access to the new publicly available statistics that will come along from the Administrative Data Research Centres? What is the problem?

Michael Fallon: There is not a problem.

Q179 Chair: Well then, why do you not just say, “Yes, they should have it”?

Michael Fallon: The network has been established to enable social science research to inform Government policy in the first instance, and that is its job.

Q180 Chair: So this Open Data Initiative is really for Government; it is not for the public or the wider economic benefit.

Michael Fallon: No, I did not mention the Open Data Initiative. I said that the Administrative Data Taskforce network in the first instance is to enable social science research to help inform Government policy using the body of information within government. I have already said—and I think Mr Hopkins described this as caution; I hoped he would have recognised that I was opening the door—that we are prepared in the future to look to see whether there is long-term potential for the business sector also to benefit from that kind of Government-held data. We will have to look at that.

Q181 Chair: The Shakespeare Review included a valuation by Deloitte of the economic advantages of Government releasing public data, and these were assessed conservatively at £1.8 billion, with wider social and economic benefits of £6.8 billion to the economy. Do you recognise the economic potential as highlighted in the Deloitte report?

Michael Fallon: Yes, I do. I think it is one of the more robust reports in this particular area. It is relatively recent; it covers information up to 2011-2012. I think it shows you the potential for significant development of this particular market. It

is a relatively new market, but direct benefits of this size are likely to have a huge impact on economic growth.

Q182 Chair: Can we take it, then, that the privatisation of the postcode address file is not going to be a precedent, and that other publicly owned address register information will be released on the open market for free to the public and to business in order to capitalise on the economic opportunities?

Michael Fallon: You cannot take it in exactly those terms, Chairman. I have said already that the postcode address file will continue to be regulated.

Chair: I am not talking about the postcode address file.

Michael Fallon: You mentioned it.

Chair: Let us talk about the other data.

Michael Fallon: Which other data?

Chair: The other data like the Ordnance Survey data and the local government address file. Those clearly are data that have been assembled by the public sector for the public. Surely they belong to the economy and to the public as a whole, not to individual businesses.

Michael Fallon: Let us have a look at the individual trading funds that I am responsible for. They are rather different. Two of those—Companies House and the Land Registry—collect data really as a by-product of their core function, which is a registration function in each case: registering companies and registering land. The other two—the Meteorological Office and the Ordnance Survey—have data collection really as their primary activity. They deliver value from collecting that data, from processing it, from maintaining it and so on, and from direct market interaction with their customers. The model we have of the trading funds is designed to allow for the delivery of revenue-generating operations by Departments in a commercially orientated manner. For organisations, particularly those like the Ordnance Survey or the Met Office that are primarily focused on the collection and distribution of data, understanding the financial costs involved and the revenue potential is critical.

Q183 Chair: So the revenue potential for Government agencies is more important than the wider economic benefits.

Michael Fallon: No, it is not more important, but it is important. There is a cost to collecting this revenue. In many cases, there are very large companies who are quite happy to pay for it, and I think it is wholly reasonable for these agencies or trading-fund bodies to charge a fee for that service. I do not think it would be right simply to give very large companies an immediate advantage in that respect. This is something we look at through the Public Data Group, which has data

responsibility for all four of the trading funds, and we keep looking at that. You will know that some of the fees involved have continued to be lowered by Companies House, for example, to make it cheaper and cheaper, and all four of the trading funds have continued to release more and more information from their datasets and to release that information for free. I can give you an example for each of the trading funds if the Committee would like it.

Q184 Kelvin Hopkins: As I understand it, for the purposes of the census, the three separate organisations that gather information on addresses were merged and we had an immensely valuable refined list of addresses that was used for the census, which was deliberately destroyed afterwards, on the instructions of Government, to protect the commercial interests of the organisations—no doubt particularly the Royal Mail. If that was kept in the public sector it could still be charged for, but it seems to me that something that could have been immensely valuable was deliberately destroyed.

Michael Fallon: I have no information on that.

Mr Hurd: I have no information on that at all. All I would say in terms of the future of the census is that very much part of our thinking is how we can make better use of existing datasets in a process that would be substantially cheaper.

Q185 Chair: So you are not aware of the fact that the census was based on a combination of three address datasets and that it was deliberately forgotten about immediately after the census, despite the fact that it had probably created the most valuable address file ever in existence in the United Kingdom?

Mr Hurd: I am not aware of it.

Chair: You may wish to write to us about the matter, in the traditional form of these things.

Q186 Mr Turner: What action is the Government taking to ensure the wider economic benefits of open data are maximised?

Mr Hurd: Our starting point is to get the data out so that we get as much of this raw material out there as possible, in a portal that works and that people find easy to use. Then, as a Government, we set up the Open Data Institute to be a hotbed of innovation and support for entrepreneurs who want to start developing new applications to exploit that data. In addition to that, there is a process of engagement with business, both large and small, through the Transparency Board, through the ODI and through our outreach programme, to raise the profile of this. There are elements of the commercial world that get it and want it and know what to do with it, and specialist businesses that exist to aggregate this data and serve it up in a useful way to other companies; there are other stretches of the economy that

are only just getting their heads around it. Again, we are at the start of the journey. I know from my own experience with responsibility for civil society that our charities and social enterprises are very much at the start of the learning process in thinking about how they can make use of the open data opportunity. Part of our responsibility at the Cabinet Office is to raise the profile of the opportunity, and we do that in large part through the kinds of organisations and institutions I have talked about, not least the ODI.

Michael Fallon: BIS does quite a lot of work in this field. We have a range of different initiatives to help companies develop new commercial opportunities. We have the Open Data Challenge Series that supports companies that want to develop new products and services using open data; we have innovation vouchers from the Technology Strategy Board that businesses can apply for to work with new suppliers to develop new business ideas that use open data; we have the Breakthrough Fund, which helps agencies and local authorities, particularly where there are some short-term barriers—for example, inadequate software—that stop them using open data. The Technology Strategy Board has also funded the Connected Digital Economy Catapult, which helps bring innovative businesses together with researchers and academia to help them over the early-stage issue to get launched as businesses.

Q187 Mr Turner: What are the two or three pieces of your work—presumably with companies in the private sector—that pull this ability in the most?

Michael Fallon: We are making money available through our various funding streams. The Technology Strategy Board is the main one, but the Catapult centres are there as well. We are making money available; we are encouraging collaboration; and we are pushing the innovative companies in this field closer to academia, which is one of the principles behind the Catapult centres, to make sure that these opportunities are not missed and that we help promote the most creative and innovative digital businesses that we have. We have a lot of them in this country.

Q188 Mr Turner: I understand that civil servants are not particularly good at using the abilities presented by the wider open data themselves. Is that true?

Michael Fallon: The kinds of funding instruments we use tend to be competitive; they tend to be challenge funds that people bid into, to ensure that we get the highest-quality bid and that we are spending our money in the best areas, so we are not really reliant on the skills of civil servants. We have externals on the Technology Strategy Board and we take external advice on the various competitions that we run.

Mr Hurd: I do not know if the Committee has visited the Open Data Institute—I know it took evidence from Sir Nigel—but that is an institute that is set up to help with this interface with business, not least in incubating start-up companies. Some

of these can be enormously valuable. One of them is Mastodon C, which developed a project called Prescribing Analytics, which analysed prescription data and identified a potential saving of £200 million per year to the NHS if GPs changed their prescribing habits. This is one incubated enterprise inside the Open Data Institute that appears to be throwing a window on hundreds of millions of pounds of savings in the NHS just by the ability to analyse data. That is just one business.

Q189 Chair: Is this free data?

Mr Hurd: I believe so.

Michael Fallon: I think it is data from the Prescription Pricing Authority.

Q190 Chair: Is the anonymised health data free to the pharmaceutical companies as well?

Michael Fallon: That I do not know.

Mr Hurd: I do not know.

Q191 Chair: Is it not obvious that the anonymised medical data should be made freely available to anybody who can use it? It cannot be a breach of data protection, but it puts into thousands of researchers' hands the possibility to develop and innovate for the public benefit with this data. That is why we are in favour of open data. Is it the Government's policy to do that?

Mr Hurd: All I am trying to single out is that, just in that one business, through data that is available, there is a very big potential upside to the taxpayer simply from pointing out savings if GPs prescribe cheaper generic statin drugs. That has opened up a debate around that, which we would not necessarily have had if that organisation had not analysed the data in that way. All I am trying to signal, Chair, is that through an incubator mechanism and through an institute that we have set up to stimulate this entrepreneurial approach to data—which, again, is outside the skill set of the Civil Service—we are already getting some interesting results. It is still very early days.

Q192 Chair: Mr Fallon, on the one hand the Government is charging, or maintaining the right to charge, for data; on the other hand it is handing out money to promote innovation in the private sector. Would it not just be simpler to let the private sector have the data for free? Would it not be more productive? Would you not be more likely to achieve what Deloitte says is achievable if we gave them the data for free? Would that not be a really big reduced barrier for business?

Michael Fallon: We make a huge amount of data available for free. I can give you examples from Companies House, which provides free access to all registered

companies that file statutory accounts. There are 3 million companies on the register; over half file their accounts digitally. All that information is free. The Land Registry make available monthly residential property price data going to go back to 2009 and they are going to go back to 1995. That is already 17 million records. There is a huge amount of information we are already putting out. However, there is some very up-to-date information that has a cost in collection that large companies are perfectly content to pay for because it has been collected by authoritative agencies, and I think it right that the taxpayer should see some reimbursement for those costs.

Q193Chair: What would you say to people who say that the taxpayer is cutting off its nose to spite its own face, because the taxpayer would do better in the long run if it let this data free to “fructify in the pockets of the people”, to quote Gladstone, rather than to be Scroogely let out at a charge to the companies that are prepared to pay for it?

Michael Fallon: The real answer is to keep it under review all the time and to see what more can be put out all the time, but, where you are going to put data out, to make sure that it is first available for the very smallest companies—for the SMEs, the innovative, the charities or whatever. There is a danger in putting data out that it is immediately swallowed up by the big global companies and the benefit is not to the UK economy but to the Microsofts, the Googles or whatever. I am a bit wary about data that the taxpayer has paid to collect and is still paying to maintain being packaged up and delivered free of charge to very large companies that are not based here that can then exploit it.

Q194 Chair: But how is it encouraging a small company to grow if the message is, “You can have this data free now but, if you grow, you will get clobbered for it”?

Michael Fallon: You are not going to be clobbered. The agencies that do charge for data have scales of charges, and these are charges that the market is bearing quite comfortably. I have seen very little evidence that SMEs’ growth, for example, is being restricted by the amount of money they are being charged for access to the postcode address file, an issue I have been looking at very closely, and I have seen very little evidence elsewhere, in being responsible for SMEs, that charges are disproportionate. But where these agencies can recoup the cost of collection, it is right for taxpayer value that they should be charged.

Q195 Mr Turner: This is the first I have heard of this deal between small and what I would call micro-businesses. Am I correct in thinking you are helping the micro-businesses?

Michael Fallon: I am saying that if we are going to start putting data out there, that is the end we should start at. We have done that with the postcode address file by making it free for charities and for the very smallest businesses—the ones, for example, who would otherwise incur the repeated cost of making multiple applications in a particular Yellow Pages area that they wanted to research to set up

a business, and so on. If we are going to continue down this route, then the way to start to help business is to help the new and the smaller businesses. All I am pointing out to the Committee is that there are large companies that are happily paying for this data at the moment.

Q196 Mr Turner: Right. But there may be, or there will be, or there is already, the need for your help from small and micro-businesses. What I am trying to work out is whether this is happening now, whether it is to happen, or whether it is an ambition.

Michael Fallon: It is beginning to happen now. The postcode address file is one example, but it is beginning to happen now. We are seeing charges that are being reduced; charges that are more proportionate; services that are being marketed much more than they ever used to be before to the SME sector, and so on. This is the end of the economy that you need to start with before you hand huge chunks of taxpayer-funded data over to large multinationals.

Q197 Mr Turner: And you could let us know which companies were benefiting from those.

Michael Fallon: I am very happy to write to you and give you more detail on this.

Q198 Paul Flynn: I find myself in strong agreement with you on the question of not handing over valuable data freely to multinationals, because they behave pretty ruthlessly. One of the pharmaceutical companies has just been fined £3 billion for its work. This is not a question of handing over scientific material that can be used to enhance their profits. It is of value to the taxpayer; we should certainly charge for it. It has been suggested by a number of witnesses that the Government is using the excuse of privacy to deny releases of information, and that progress towards outsourcing means a smaller area in which information is available. Tom Steinberg told the Committee—he has got a distinguished record of putting information into the public domain—that, “Our children will know much less about how the Government works and whether or not it is doing a good job than we do”. Is that fair?

Michael Fallon: No, I am afraid I do not think it is. We know far more about the internal process of Government now than we did 20 or 30 years ago.

Paul Flynn: Indeed, but he is talking about the next 20 or 30 years.

Michael Fallon: I suspect our children will know even more, because there is a process under way here—and I think your Committee supports it—that more data should be made available, and it is the Cabinet Office that is the one encouraging Departments across Whitehall to make it available. There are always going to be issues of privacy in the open data area; issues of personal information; issues of

commercial confidentiality of information that is supplied by companies to Government for one purpose; and issues of national security.

Mr Hurd: I agree with Michael on that. I agree that we have much more information than we have ever had about what Government is doing in our name, but I also think this process is now irreversible, not least because of how the public are reacting to what is being put out. I am not just talking about the reaction to the crime maps, but if you look at the new platforms that people have set up to provide the ability to follow the money and drill down into where the DfID budget goes; the analysis on local government spending; and the ability to compare through things like the Birmingham Civic Dashboard, the public are responding to this.

I cannot really imagine a situation where any Government of any colour in the future could put its hand up and say, “You know all that data and information that the other lot gave you? We are going to give it back and shut it up again”. I just cannot see that happening, but it does make it even more important for us to establish this open-by-default approach and push this process as hard as possible so that we get the data out, the entrepreneurs and the developers grab it and use it and play it back to us in new ways, and we say, “Oh, that is valuable”. We have not talked about the transport area, but the Shakespeare Review was very heavy on that. Our whole experience of interfacing with our transport networks—the buses and things—has been completely changed by all this. People do not have to wait for a bus anymore. The more this becomes alive and relevant to people, the more people are going to want it and are going to reject the suggestion that we go back. Now the door has been opened, it is going to be very hard to close it again. I, for one, welcome that.

Q199 Paul Flynn: In this golden age of openness and transparency that you have created, a national newspaper recently asked for the release of certain documents. When they were refused, they put a freedom of information request in. That was acceded to. The Government again refused and it went to court. The court found in favour of publishing these documents in the public interest, yet the Government are still refusing to publish them, on the ground that if they were published, they might show that the heir to the throne is not fit to be King. If such information is available, should we, as his future subjects, not know about it and possibly make other arrangements for an alternative heir to the throne?

Mr Hurd: I am not going to comment on that, but I would like to challenge your opening assertion that somehow we are in a golden age. We are not. That would sound complacent, which we are not. We are on a journey, started by the last Government with the Freedom of Information Act and built on by this Government, to make government much more transparent and accountable. I do not think anyone is claiming that we have completed that journey; I think most of us feel that we are still at the start of it. We are very far from a golden age. What I am concerned about is that we make enough progress for this to be irreversible, because it is definitely the right direction of travel.

Q200 Paul Flynn: Why should the public be denied knowledge of what probably the most influential lobbyist in the country is saying to Government in his letters and in the 53 meetings that he has had with your Government since 2010? Why on earth should we not know what is going on? Why do you deny freedom of information requests and a decision by the court in order to retain this information?

Mr Hurd: I am not here to make any comment on specific situations where a request has been denied.

Paul Flynn: Why not?

Mr Hurd: Because the system is set up whereby Government has the right to make a case for denying a request. I do not know enough about the individual situation to comment on it, but that is a process whereby Government can continue to make a case for saying, “No, we do not think that is appropriate”.

Q201 Paul Flynn: The suggestion was made that, if published, those letters would show that the heir to the throne is not politically neutral, and that is the main qualification for an heir to the throne in this country because they have to take decisions. If a Prime Minister is acting in her or his own interests and not the nation’s interests, the Head of State has to intervene. One would suggest that the Head of State is possibly not the right one if he has views that are not are politically neutral. What is the case for keeping that secret from us?

Mr Hurd: I am not part of that process, Mr Flynn, so I am not going to comment on it. All I would say is that under the Freedom of Information Act, there will always be situations like that—it happened under the last Government and it will continue to happen under this Government—where people will disagree with those kinds of judgments. My point is that we are in a much better place than we were. The fact that we can even have this discussion would not have been possible 15 years ago.

Q202 Paul Flynn: Does this not suggest that the Government is infantilising the nation by telling them, “This information is there, but you should not know about it because it might confuse you”?

Mr Hurd: No. The last thing we are doing is infantilising the nation. What we are saying is: “We are sitting on a huge amount of data. Many of you people out there have got a better idea of how it can be turned to public value. We are going to get it out.” We are going to change the culture of government so that we get out all this stuff that we have been sitting on for years, which we probably do not know the value of. That process is going to be a bit clumsy and slow; it is going to frustrate a lot of people, like the witnesses that you have had who want us to go further and faster, but it is important to recognise how much progress has been made and that this country is leading a global process whereby other countries are embracing this

as well. That is also in the national interest, because that generates commercial opportunities for our businesses to exploit data that has been put out by other countries.

Q203 Paul Flynn: You would not share the view of the Prime Minister who took us into the Iraq War that the worst thing he did was the Freedom of Information Bill; you think possibly that it was an advance.

Mr Hurd: I found myself on quite regular occasions not being in agreement with the previous Prime Minister. That is my default setting.

Q204 Chair: Iraq is a common theme of this Committee's deliberations. Mr Flynn, thank you very much.

Thank you very much for your presence with us this afternoon. Before we let you go, can I just return to the question about the relationship between the requirements of privacy and the principle of open data in respect of medical information? This may be something you want to write to us about rather than answer now, but Mr Shakespeare, in his submission, said that public bodies should “stop being hyper-nervous about privacy risks of open data to the point where we obstruct, for example, life-saving and life-enhancing medical advances”. He goes on to say that this has meant “a policy that case-level medical data should not be shared with experts conducting research for commercial companies, but the greatest potential lies in the cross-NHS case level data. Making that available only to academic or Government researchers would mean huge opportunities would be lost.” Would either of you like to comment?

Mr Hurd: I am very happy to write to you on the specifics. The issue touches on some fundamental potential tensions in this process, which we think are eminently manageable. Our position is that data should be open by default, but subject to important caveats, including personal privacy. A number of your witnesses recognised that part of the cultural challenge here is what went on before, with data being lost and tremendous sensitivities around that, but the truth is that anonymisation techniques mean that data can still be released while providing protection to individual citizens. We need to be careful to make sure we are up to speed on the technology there, but those techniques apply. There is some personal data that will never be made open data, though it can be appropriate to share this data across Government or with trusted third parties as long as strict processes and legal frameworks are followed. Our view at the Cabinet Office is that these risks are manageable and we wish to actively manage them. I am more than happy to drop you and the Committee a note on the specifics you are talking about.

Q205 Chair: What risk exists with anonymised case-level data?

Mr Hurd: That is why I said that through anonymisation techniques—

Chair: Does your answer not underline what Stephan Shakespeare is saying, that the reflex is to be ultra-cautious?

Mr Hurd: It is. Anonymisation techniques are there, but, again, we need to be careful. There was a case in point in 2006, when AOL released a batch of supposedly anonymised internet search history data for research purposes and bloggers and journalists immediately focused on clues to identity and successfully triangulated that with some very embarrassing search queries; AOL were swiftly forced to withdraw the programme. In a field where the technology moves on and there are lots of very smart people out there, we need to be quite sure that we are absolutely at the front of the curve on that. Anonymisation techniques certainly can mean data can still be released while providing protection to the individual, but we do need to be clear that there are still some risks attached to that. The fundamental point is that if we are to take the public with us, they need to trust the process. If we lose that trust, then this agenda will crumble.

Q206 Chair: What do you think is the difference between the generations? Older generations seem to be more cautious about data release, but younger generations seem happy to splurge their personal data all over the internet.

Mr Hurd: They are for the moment, Chair. As a father of teenagers myself, I caution them against that. I am not sure that we are at the end judgment on that in terms of the attitudes of the generation coming through, because I do not think they have yet been confronted with evidence of the downsides attached to that.

Q207 Chair: So you do not think there is a difference between the generations.

Mr Hurd: No, I absolutely take your point. I am just not sure that we can make final judgments on that quite yet, because we are still at a very early stage in that process of personal openness, as it were.

Q208 Chair: I am grateful to you both. I am bound to say that I think many of our previous witnesses would summarise the Government's position as being very willing in principle but finding it very hard in practice to be enthusiastic about open data. It is likely that our findings in our report are going to reflect that. Is there anything you would like to add, just to give you a right of pre-emptive reply, if that were our conclusion—that the rhetoric is somewhat ahead of what the Government is actually prepared to do, particularly in regard of trading funds?

Mr Hurd: If I can be blunt, the Committee can focus on trading funds and on a specific instance where a debate was had and a decision was taken, but I would urge the Committee not to lose sight of the thousands of datasets that are out there; the enormous progress made in terms of getting data out, changing the culture of Government to publish, and upgrading data.gov.uk so that it is recognised as the

leading data portal in the world; the architecture that we have set up of what I call challenge and advocacy, so that we have exactly the kinds of witnesses that you have had in front of you actually sitting in government giving us a harder time; the creation of institutions like ODI that are helping us engage with the people who really understand data and see the potential; the enormous step taken in publishing an inventory, effectively, of all our datasets; and the fact—recognised around the world, including at OGP, which is a meeting of 60 countries around the world, and recognised by all the recent independent surveys—that this country is leading the world on this critical agenda. We are not remotely complacent about it; in fact, we are very alive to all the difficulties and challenges, but we think—and it is recognised by other people—we have made enormous progress in terms of the reality.

Q209 Chair: Thank you, Mr Hurd. Mr Fallon, do you have anything to add?

Michael Fallon: Yes. I hope you will not rush to a verdict too quickly, and I would welcome the opportunity to write to you about the trading funds and the work that is going on in the Public Data Group, the way in which each of those trading funds has approached the release of data, the huge amount of data they have already got out there, and what the policy is in terms of fee-charging and so on with each of them. Perhaps it would be fairer if I did set that out fund by fund.

Chair: That would be very useful. Thank you very much. I am most grateful to you both.

Examination of Witnesses

Witnesses: **Jil Matheson**, National Statistician, UK Statistics Authority, **Glen Watson**, Director General, Office for National Statistics, and **Sean Whellams**, Head of Statistics Profession, HMRC, gave evidence.

Q201 Chair: I wonder if each of our three witnesses in this second session on open data could identify yourselves for the record, please.

Jil Matheson: Jil Matheson, National Statistician.

Glen Watson: Glen Watson, Director General of the ONS.

Sean Whellams: Sean Whellams, the Head of Profession, Statistics, in HM Revenue and Customs.

Q211 Chair: I am most grateful that you are with us. Ms Matheson, we are here to encourage you to take on your more public role that we recommended in one of our recent reports. Therefore, I am looking forward to you answering our questions. First of all, Ms Matheson, can you share with us the most significant new ONS statistical series that has been published as a result of the Open Data Initiative?

Jil Matheson: I am very happy to talk about not just the ONS but the Government Statistical Service as a whole, because the GSS is one of the primary sources of data that appears on data.gov.uk. The single biggest source of data there is from ONS, but over half of all the data that are released on that portal are provided from the Government Statistical Service. That reflects a longstanding passion not just of mine but of the statistical community to get our statistics and our data used, and used in useful ways. There is a history, going back, of providing survey data to the data archive at Essex University in anonymised forms, right through to what we do now, which is providing data in different formats and on different platforms for citizen users—that is the jargon for people who want headline statistics, perhaps on a mobile device—through to expert users who want data to download into their systems, and providing access in secure environments to data for academic researchers. That has been part of what we have been trying to do. What the Open Data Initiative has done is to give that an impetus, to give it a set of expectations, to shift the culture and to encourage that. Glen will be able to talk about the ONS in particular, but statisticians in Departments have responded to that with alacrity.

Q212 Chair: Can you give us an example of a new series that is now being published as a result of the Open Data Initiative?

Jil Matheson: There are series that have been published in different forms. The best example probably is the census. Again, Glen can talk about how it is not just about what is published—it is about what is published—but also about the forms in which the data are published.

Q213 Chair: But there must be a lot of fresh data that we are publishing that we have never published before.

Jil Matheson: As I said, there is a long tradition of data underpinning statistical series being published not always in useful and machine-readable form. It has often been locked up in paper publications or in PDFs. There are still a lot of PDFs around on websites.

Q214 Chair: In terms of top-level data, there is nothing new being published.

Jil Matheson: The agenda changes all the time; we are always publishing new data. The Open Data Initiative itself is about encouraging reuse of existing data; it is not necessarily about providing new data. Let me give you an example. ONS

publishes every month the Consumer Price Index, the Retail Price Index and other price indices. It has done so going back to the 1940s. What we now do is publish the underlying data in a form that can be used and reused. It is much more finely aggregated than was previously available. That is a result of the technology and open data. There are other examples like that. The headline statistics were there, but the underlying data in forms that are reusable is what is new.

Chair: Mr Watson, do you have anything to add? Perhaps you have got an example.

Glen Watson: A good example would be the publication of all of the underlying ad hoc data requests that we receive on the website, which we introduced 18 months or so ago. We answer about 2,000 ad hoc requests for data for analysis a year. Prior to the open data movement, we were answering these questions without publishing the results. We now publish the results on the website for all of them, or, where we cannot, because of volumes or whatever, we publish a list of those that have been answered and say that the data is available on request. That is one practical example.

Chair: Mr Whellams?

Sean Whellams: Only last month we produced a breakdown of our tax receipts, which we produce on a monthly basis, but this time we produced an extra series, which broke down those tax receipts between the four countries, England, Northern Ireland, Scotland and Wales. That is a brand new source that was released last month.

Q215 Chair: We will be very interested in that during the devolution and separatism debate.

Ms Matheson, what do you see as the role of your statisticians in making open data more of a reality?

Jil Matheson: They have got a hugely important role. It is partly about getting the data out there. It is about encouraging use. It is about standards. For us, this is about making data useful as well as making it available. It is not just about publishing it; it is about publishing it in ways that people can understand. It is about standards. It is about that technical phrase “metadata”; a number does not mean anything on its own. It is about being able to describe the number and know what it represents. It is about skills and capability. For example, one of the things that the Government Statistical Service has done is produce an e-learning module that is available to all civil servants about what open data is, what the standards are and how it should be applied. It is about signposting and pointing users to the data that are available. Again, they are not going to be used if people do not know they are there and cannot access them. It is about the context. Whenever we publish something—I know this from many years’ experience—the first question we get asked is, “Has that changed over time?” or, “How does this compare with another part of the country?” or, increasingly, “How does this compare with another country?” That contextual, comparative information, the authoritative, impartial reporting of that data and the skills that go into processing data, quality-assuring it

and providing privacy protection, so that the data that we publish are anonymised, not disclosive, and protect the privacy of the individuals and the business whose data we hold, are all part of the role that statisticians have. They are not alone. One of the phrases that I use is, “This is a team sport”. It brings together lots of people with lots of different skills. Statisticians are a crucial element in that team.

Q216 Chair: What do we need to do to make sure that people advising Ministers on policy understand all this newly available data and make best use of it?

Jil Matheson: Again, one of the things that I personally have been doing is working with the policy profession and the head of the policy profession across Whitehall to make sure that this agenda is part of the skill set not just of statisticians or analysts but of policy professionals too, so that they know what questions to ask and they know that they have skills in their Departments.

Q217 Chair: Apart from pleading with the policy profession, what are you actually asking them to do?

Jil Matheson: It is more than pleading. There is an important role for Ministers, and non-executives in particular, as well as other bodies, in saying, “Okay, this data is there, but use it”. One of the things I observe is that Departments themselves do not always use the data that they have, let alone the data that is more widely available, to drive their business decisions.

Q218 Chair: What training do the policy professionals have in the question of statistics and data?

Jil Matheson: Until fairly recently, very little. Again, part of the work on Civil Service reform and the Capabilities Plan is recognising that they need to, as I said, be able to know what questions to ask, how to interrogate data and how to find what data is there. That is part of the development of capabilities; the GSS provides that training. It is not just providing training for other civil servants; we have a role in going beyond that to stimulate and support the users outside and the data-literacy among the general population.

Q219 Chair: As UK National Statistician, Head of the Government Statistical Service and Chief Executive of the UK Statistics Authority, that role including executive responsibility for the ONS, do you find yourself dealing with conflicts between these different roles?

Jil Matheson: These are all completely aligned. If you look at the Statistics and Registration Service Act, it talks very early on about the role of the Statistics Authority, which is to promote statistics data for the public good. That is what we are talking about here. It is part of the Government Statistical Service strategy and mission to encourage use and reuse of statistics. A fundamental part of the code of practice against which National Statistics are judged is their openness, and it is

central to the ONS strategy too. All of these things complement each other. There is no conflict or tension.

Q220 Chair: Surely there are times when you will be advising that certain statistics should be used in a certain way, or should be released publicly, and there will be political pressures from those who say they do not want those statistics released in that particular manner, or at all.

Jil Matheson: An integral part of the code of practice is that statistics are released. That is what the assessment function of the UK Statistics Authority is there to judge: that if data are collected for statistical purposes, then the statistics should be made available. Again, it is part of the mission of the public good.

Q221 Chair: So you do not find yourself having to deal with any conflicts.

Jil Matheson: In this area, no.

Q222 Chair: In what areas do you deal with conflicts?

Jil Matheson: We have talked about this before. Inevitably—you would expect this—there will be findings from the assessment function that are not always comfortable for producers, and I will take a view. Sometimes that might mean me saying, “I do not think that is right” or, “I do not think that is important”. But in this area of open data publication, there is no conflict at all.

Q223 Lindsay Roy: You mentioned public good. Can I ask you what test you use for publication of official statistics of the public good? What are the benchmarks?

Jil Matheson: The criterion is that if this is data collected for a statistical purpose, it shall be published. The only qualification on that is that that is always subject to privacy protection. We will not release data if it is likely to identify individuals, individual households or individual businesses. That is part of our mission. It is part of our legislation, but it is also part of our code of practice and our culture.

Lindsay Roy: So if you cannot anonymise?

Jil Matheson: Anonymisation is more than just taking off names and addresses. If it is likely to be disclosive then we will do one of several things to try to protect the data. Often, you can release it at a higher geographical level or a higher level of aggregation and the problem of disclosure disappears; or there are techniques that you can use to protect the data at smaller level of aggregation; or you can provide access to anonymised data in safe settings, or safe havens, so that researchers or users never directly have access to identifiable records. There is a whole range of

things that you can do, but there is a fundamental requirement and desire to protect privacy and confidentiality; that is the basis on which we operate.

Q224 Lindsay Roy: Have you found that to be secure?

Jil Matheson: There have been no leaks of that. There are a lot of very sophisticated techniques—statistical techniques, but also technology, legal protection and culture. There is a whole suite of things you have to have in place in order to be able to do that. I know some of our users would like us to be less cautious. I am aware of that, but I am also aware of our obligation to protect privacy and confidentiality. I had a discussion very recently with a senior executive of a leading IT company who was very keen to work with us to develop means of access to data, and halfway through the conversation he said, “But you will keep my census data secure, will you not?” Of course, yes. It is about finding the way in which we can protect privacy whilst maximising use, utility and access.

Q225 Kelvin Hopkins: We pressed the two Ministers we saw just before your good selves on the performance of Government Departments, and they were very cautious and measured in what they said; they were not really very forthcoming and they did not want to be critical. We might get a more objective view from you. From your point of view, which Departments are the most and least helpful to the open data agenda?

Jil Matheson: I did see some of that, and I was delighted to see that ONS seemed to come out as the most helpful. That is good. I agree with that. The least helpful is really difficult to say, because every Department that I know of releases its statistical data, subject to the kinds of controls that we were just talking about. They do not all do it in a very timely way. They do not all do it in a completely open way—i.e. the formats are not always completely open; Departments will be at different places in that. But every Department that produces statistics has data on data.gov.uk.

Q226 Kelvin Hopkins: You are replying that some are more cautious and tardy than others. Would you like to compare the Departments? Which is the most enthusiastic and wants to give it to you all the time and is quite open, and which is the one that fails to do so?

Jil Matheson: It is easier to talk about the most enthusiastic, because they are the ones that I see, and I do not get complaints about the others, apart from ONS. One of the reasons that we have just recently published a data strategy for the Government Statistical Service—Sean has been leading the work on that—is to encourage everybody to be as good as the best. Some of the good examples I could point to would be DWP, with both the way that it handles the ad hoc requests—the kind of thing that Glen was talking about—and its means of access. It has a tool

called Stat-Xplore—again, Sean would be able to talk about that. Some of the things that the Ministry of Justice has done have been really encouraging. Some of the things, I have to say, that HMRC has done have been excellent as well. It is about learning from those, sharing what is really good and encouraging others to come up to those standards.

One other point I would make is that one of the issues for statisticians is that data.gov.uk is the mechanism by which most Departments release their data, and it does not yet have the functionality that we would like to see for accessing statistics. One of the really important ways of people being able to understand what is there is to be able to visualise it.

Q227 Kelvin Hopkins: I dwelt on the historic culture of secrecy in British Government, which is almost legendary—it certainly was the case when I was a student a long time ago. Would an annual audit of the open data performance of each public body be a good idea?

Jil Matheson: The publication plans and the National Information Infrastructure should make that possible, with commitments from Departments as to the data that they will be releasing. They will be able to be quizzed on that and their performance assessed.

Q228 Kelvin Hopkins: We have league tables for schools; would a league table for Government Departments concentrate a few minds, do you think?

Jil Matheson: I am not a huge fan of league tables for schools. It does depend on how much data they have got, which is not equal across Departments. There is a cost and a resource to doing this too, and some data are going to be of much greater interest to a wider range of users than others.

Q229 Kelvin Hopkins: Could you be more proactive when it comes to publishing new GSS statistics as a result of newly-available data? Do you have problems getting data from some Departments?

Jil Matheson: One of my frustrations is that I know there are administrative data in some Departments that would be fantastic for statistical purposes in another Department, including in ONS, and it is extremely difficult, in many cases, to get access to that data.

Q230 Chair: Can you give us an example?

Jil Matheson: I thought you might ask that, and I wanted to give Mr Whellams a bit of notice, because one of the examples is the data in HMRC. For example, data

on corporation tax would be extremely useful for ONS in production of national accounts and other economic data.

Q231 Chair: Sorry, you will have to explain that a bit more to the layman. We know how much corporation tax is collected.

Jil Matheson: We do not know at a company level.

Q Chair: What, how much each company is paying?

Jil Matheson: How much, yes.

Q232 Chair: Do they not tell you that in their accounts?

Jil Matheson: Not in a very timely way. You need to link it with other kinds of data to be able to say what the activity of the company is, what it does and what sector it is in, because that is how the accounts are broken down.

Q233 Chair: But the data eventually becomes available through Companies House.

Jil Matheson: A long time afterwards. What we do at the moment is send out forms to businesses.

Chair: And they fill them in.

Jil Matheson: And they fill them in. There is a cost to ONS and there is a cost to the business.

Q234 Chair: What are the issues involved that you feel need to be addressed before HMRC can hand over that administrative data?

Sean Whellams: We do not do it currently for corporation tax data, but we have got examples where we do it for the Inter-Departmental Business Register, which the ONS pull together. That is based on VAT data and pay-as-you-earn data. In order for that to have happened, primary legislation would have needed to go through, in order for HMRC to be legally able to pass that data on to another party, because we are constrained by the Commissioners for Revenue and Customs Act with regards to giving taxpayer-disclosive data to—

Q235 Chair: Is that a universally held view of the law?

Sean Whellams: I believe so, yes.

Chair: So nobody disputes the fact that you are prevented from handing over that data because you are not empowered to.

Sean Whellams: Absolutely.

Chair: So we need legislation.

Sean Whellams: In order for HMRC to pass taxpayer data to the ONS or another party, yes, we would need primary legislation.

Q236 Kelvin Hopkins: I am glad you mentioned HMRC. I would love to spend the next half an hour talking about HMRC, because a serious problem is tax evasion, tax avoidance and the under-resourcing of our tax collectors. I understand from some of the trade unions, for example, that every tax collector collects many times their own salary, so logically, if we just employ many more tax inspectors, we will start to solve the problem. Government has been reluctant to do this. I understand that HMRC has been under pressure not to be—

Chair: Do you have a question, Mr Hopkins?

Kelvin Hopkins: This is statistical data that could be made available to you that could help politically to drive one of the most important things we have, which is to disclose the tax gap. That sort of data is just what we want. Are you able to provide it?

Sean Whellams: We provide an annual publication, called *Measuring Tax Gaps*, which is a really thick publication that shows a large amount of information about the difference between the tax we have collected and our estimates about how much we should be collecting. We do that, and that is cut in many different ways. In the last Spending Review, the analysts within HMRC pulled together some analysis that was supporting HMRC's business case for more funding, and HMRC did get some reinvestment money. That was based on the central analytical advice, which was then scrutinised by Treasury, in order for us to get a larger settlement in order to put tax people on the ground and collect more tax.

Q237 Kelvin Hopkins: I could go on a lot longer on that point, but I will not; I will leave it there.

The ONS spent quite a lot of money, I understand, on creating the census register, using three major sources to provide a very effective register, and it found that 1 million-plus addresses were missing from the postcode address file. What happened after that was that the Government refused to allow that refined and better address register to be used, and it has effectively been destroyed to protect commercial interests—especially those of the Royal Mail, but those of privatised or Government agencies—to make money. Is that not really against the public interest? It would definitely be in the public interest to have that census register available.

Jil Matheson: I will ask Glen, who was the census director, to talk about that, but stepping back, it is fundamental to effective statistics and to open data—and lots of other purposes too—that there is a high-quality, widely used, available and accessible address register. It is the thing that links together, or makes possible linking together, lots of other data. As you said, in the lead-up to the 2011 census, it was clear that we did not have a comprehensive, useful and consistent national address register, so ONS did a lot of work to quality-assure. Since then, GeoPlace has been set up. ONS works closely with GeoPlace. As far as I know—again, Glen will be able to add to this—the quality of the address register as it currently exists is maintained, so it is comparable with the quality that we helped to develop in 2011, for most addresses. We have some concerns about institutions—communal establishments—that are not very well identified or listed or specified on the address register, so there is still a lot of work to be done. But the basic address register currently is adequate for the kinds of purposes that we are interested in. You mentioned the postcode address file, which is a fundamental part of the address register—it feeds into the address register. We want to make sure that the quality of that is maintained, and that it is accessible, available and, again, widely used for statistical purposes.

Glen, do you want to add anything?

Glen Watson: I do not have very much to add. The work that we did to create the census address register was a great catalyst for getting the address players—Royal Mail, Ordnance Survey and local government—to work collaboratively together, and the new addresses we found and the mismatches that we resolved were fed back into those independent address sources and the quality of them all took a big step forwards as a result. As Jil has said, with the work of GeoPlace since, all of the methodologies that we developed, all of the new addresses we found, etc. have fed into this ongoing process of creating a high-quality address register. It is fundamental not just for the census or any alternative to a census in the future; it is fundamental as a sampling frame for household surveys and it is fundamental for the emergency services and many other different uses across the public sector, so we strongly support the need for a good-quality address list.

Q238 Chair: Will the public sector now have to pay Royal Mail for access to the postcode address file?

Glen Watson: My understanding is that there is a public service mapping agreement that is co-ordinated and paid centrally by the Department for Business Innovation and Skills.

Chair: But there is a payment.

Glen Watson: There is a payment associated with that, in order to be able to acquire the necessary address information from Ordnance Survey.

Q239 Chair: Did the Government Statistical Service make representations before the postcode address file was included in the Royal Mail sale?

Jil Matheson: No, we did not.

Q240 Chair: Did the UK Statistics Authority?

Jil Matheson: Our interest is to make sure that the PAF is maintained and is available and accessible to us, of course, but, beyond that, to all users. There was a clause in the Bill in 2011 that talked about making it available to everybody on reasonable terms. We will keep an eye on making sure that those terms are reasonable.

Q241 Chair: What residual concerns might you have about the accessibility and easy availability of that public information?

Jil Matheson: At the moment, we have no concerns about the availability for use within Government. What we want is for the postcode address file and the address register to be available to anybody who is providing data. The risk is that if the PAF is not accessible and freely available, or if the address register is not accessible and freely available, people will stop using them and we will end up in the position that we were in pre-2011, when we had lots of different competing and not-very-good sources that people had developed for their own particular purposes but that did not talk to each other or were not consistent with each other. That would be a detrimental step.

Q242 Kelvin Hopkins: It strikes me that you have done something extremely valuable: you have compiled the best address register we have ever had for the purpose of the census, but that has now been either locked away in a cupboard or destroyed. It could have been an enormous resource for the public sector. There is no reason why the public sector could not charge for this information.

Jil Matheson: As we have tried to explain, it has not been. All the work that ONS put into developing that register for 2011 has led to a real improvement in the underlying sources that are currently being maintained by GeoPlace. We do not have current concerns about the quality of the address register, apart from in the one area that I mentioned, which was communal establishments. Ensuring that that is maintained is the critical thing.

Q243 Kelvin Hopkins: One of our Ministers referred more than once to the terms “market” and “emerging market” and the importance of promoting these markets for essentially private profit, especially in the case of the Royal Mail. That was of more interest, apparently, than having this very valuable address register compiled by yourselves as a public resource, which we could have charged for. It could have been the best register we had, in the public sector, and made money, indeed, for the public sector, which sounds like a very good idea.

Glen Watson: At the time, had we announced our intention to do that, we would not have been able to create the census address register. The only reason we were able to do so was that we could give assurances to people like Ordnance Survey and others that their intellectual property would be protected in what we did. Our job was to get the census done really effectively. I think we did the right thing, in the sense that we created the thing; we did not then sell it on, but we did lobby hard for creation of a national address register. You will remember Sir Michael Scholar writing to the Government and making the case for that quite strongly.

Kelvin Hopkins: I could go on, but I think I have made my point.

Q244 Chair: So if we do another census, we will have to go through the same exercise of borrowing the data from different people.

Glen Watson: No, because if the GeoPlace initiative continues and is successful, then we will be able to simply pick up the address register from GeoPlace and use that. For the residential addresses, we are very happy with the progress they are making; it looks like it would meet our quality standards. As Jil has mentioned, we just have some continuing concerns about the institutions—the communal establishments—and we are continuing to work with GeoPlace. GeoPlace have done the work that we had to do as a one-off on an ongoing basis, bringing together and confronting and resolving the different address sources on an ongoing basis.

Q245 Chair: Will GeoPlace be freely available to the public?

Glen Watson: It will be freely available to the public sector once the public service mapping agreement payment has been made.

Chair: So it is going to be treated as another trading fund.

Glen Watson: My understanding is it will not be more freely available than that, although, as Jil is saying, that is what we would prefer to see, so that it is there as a national asset to be exploited.

Q246 Mr Turner: What is your knowledge of the quality of returns in the different geographical departments—Northern Ireland, Scotland and Wales?

Jil Matheson: Certainly traditionally, the quality of the geographic data underpinning Northern Ireland data has been good, and it still is. They have separate lists and registers; they have their own Ordnance Survey organisation, in effect, and different registers, the quality of which is good. We have been talking about addresses. There are different kinds of geographic data available too, which is supplied through ONS: digital boundaries and digital descriptions of areas, which can be freely used, and that applies to the whole country. Apart from

Northern Ireland, which has always been the exemplar of address registers, there is not a lot of difference between other parts of the country.

Q247 Lindsay Roy: Apparently you produce a commentary to go with the datasets.

Jil Matheson: Yes.

Lindsay Roy: How do you ensure that that is impartial?

Jil Matheson: It is fundamental to what we do. It has to be impartial. It has to be authoritative. The job of the chief statistician in every Department, including in ONS, is to ensure that what that commentary is doing is bringing out the key headlines in a consistent way so that you can be certain, if ONS is saying, to take an example of what we have done, that the Consumer Prices Index is whatever the number is this month, that has been calculated in exactly the same way as it was a month ago and a month before that. There is a consistency of approach. The commentary is always impartial. People would tell us if it was not, and that is not a complaint we have. The complaint is usually that we are a bit reticent in drawing out key messages and the statistical story. We are doing a lot of work across the Government Statistical Service on improving the commentary and the presentation of data to make sure that it is accessible and understood.

Q248 Lindsay Roy: So we can be assured of rigour in the monitoring and quality assurance.

Jil Matheson: Absolutely. It is part of the assessment arm of the UK Statistics Authority to make sure that that is the case.

Chair: Thank you very much indeed to the three of you. It has been very useful. We look forward to reporting on this topic. Thank you very much.

