'What have the methodologists ever done for official statistics?'

1. When I was asked to be a key note speaker at this conference we had a conversation about which area of my work I might talk about in line with the conference theme. It was agreed that the work of the Big Data team within Methodology (which I established back in 2014) would be the most exciting and engaging. When I came to think in more detail about what I might say I questioned this decision.  
   Why did we think that the work of the Big Data team would be the most exciting and engaging?  
   What about all the other excellent work that is undertaken across Methodology?  
   Why didn’t I naturally choose these other areas of work to talk about?  
   I think the answers to these questions lie in the theme that is underpinning the conference today and is something I want to explore through this presentation. Why does the work of methodologists often get overlooked or undersold and what can we do about it?
2. I therefore came up with the title of the presentation, ‘What have the methodologists ever done for official statistics?’ which is a nod to the scene from the Monty Python classic ‘The Life of Brian’, ‘what have the Romans ever done for us’. ‘The Life of Brian’ is a religious satire comedy set in biblical times. The scene focuses on Reg, played by John Cleese. He is planning a conspiracy against the Romans. He is trying to give a speech to motivate his group of activists (the Peoples Front of Judea) into action by asking ‘What have the Romans ever done for us’ – as you’ll see if doesn’t quite work our as he expects.  
     
   PLAY CLIP
3. So what I’d like to explore today is why is it that people often question the value of the work of methodologists? Why is it that our work gets overlooked or not recognized like the impact of the Romans in the sketch? I think there are 3 key reasons. In no particular order:  
   - Methodological work is often hidden   
   - Methodologists are not the most extrovert and don’t naturally shout about our work  
   - Methodologists have at times had a reputation for wanting everything to be gold plated and that we live in ivory towers  
   I’ll take each one in turn, describe the problem but also put forward solutions
4. Methodological work is often hidden or ‘under the bonnet’. Keeping the analogy going it is the engine room of official statistics. It isn’t the public facing survey or Census data collection exercise, or the high profile negotiations with other Government departments or commercial organization for access to data. Neither do methodologists usually get involved in the press release or the glossy visualization that sits alongside the output. The methodological work mostly comes in between. It is the often complex and some might say ‘unsexy’ yet technical work that sits in between and underpins the production of the output. Referring back to the clip, something like ‘sanitation’ – essential but not the most headline grabbing topic – well until it goes wrong! So the solution here is quite simple – we need to improve our communication. Explain the complex technical work in ways people can understand. We need to make sure we get the recognition we deserve when it comes to outputs and measuring impact.
5. However, this won’t always come naturally. And this brings me to my second issue that methodologists are not the most extrovert and don’t naturally shout about our work. This is something we really need to improve and at the ONS through the Garratt Review of Methodology (undertaken last year) and the creation of the new Directorate ‘Methods, Data and Research’ we are transforming Methodology. Part of this also involves improving the way we communicate our work both internally and externally and making complex work more accessible. There are times when we need to prepare detailed technical reports to ensure transparency around methods but this doesn’t stop us also writing blogs and short articles alongside final reports or as we develop new methods in order to raise awareness of our work. But a warning – it is all about quality not quantity and the impact that communication has. We shouldn’t be obsessed with writing x number of blogs every 6 months – it is more about being smart about the way we promote our work and ensure it has impact – perhaps a challenge for us all as a GSS community to think about how we do that across Government and more widely and in a coordinated way. We need to be more confident about sharing our work at earlier stages of development, releasing outputs when they are still experimental in order to get feedback and buy in early on. We need to articulate the impact and benefits of our work – for example how important the roads or the irrigation system are. ‘What have the methodologists ever done for official statistics’ – a great deal but we don’t always shout loudly enough about it.
6. This leads me onto the third key issue that I think is an often unfair criticism of methodologists. Where we are criticised for wanting everything to be gold plated and we do not deliver at pace and support the implementation of the methods. This is a tricky balance to make. I believe that methodologists should be the statistical conscience for official statistics – ensuring that methods are robust and quality standards adhered to. With increasing use of new data sources and a drive for relevant and fast paced statistics we need to make sure these aims are not met by sacrificing quality or rigour or our integrity. I don’t think we’d have wanted to Roman’s to cut corners when building their aqueducts! These are times when our statistics are under intense scrutiny – just think about all the debate around migration figures or the impact of Brexit on the economy – our statistics are integral to that debate and we therefore need to be confident in the methods that underpin them. So what should we do? At times it may be necessary to be more pragmatic about approaches but this shouldn’t be at the expense of ‘fit for purpose’ – at the heart should be the user requirements. As outlined before we need to be more confident at sharing early work, adopting agile approaches to demonstrate progress and focusing on delivery and working collaboratively to ensure methods are implemented. If we are successful in working collaboratively with internal colleagues, e.g. topic experts, IT developers etc and external partners such as academic, commercial organizations then these teams should collectively work together to understand what is ‘fit for purpose’. And this is something that again we are addressing through the transformation of Methodology at ONS.
7. So one of the key things we need to do is to better communicate and promote our methodological work. I’d like to spend some time doing just that, trying to re-address the balance - highlighting some examples of the excellent work that has been delivered by methodologists that have made a real difference to official statistics. Like the medicine, education and definitely the wine! But I’d also like to identify some of the challenges that will require methodological input in order to support the future vision for official statistics.
8. Examples:  
     
   - My first example relates to online data collection. We have long standing expertise in questionnaire and survey design. Recent projects have involved re-designing the questionnaires used by the Monthly Wages and Salaries Survey and also the UK Innovation Survey in order to transition these surveys online. Methodologists have also designed a pilot sample of the Monthly Wages and Salaries Survey to allow mode effects to be tested without interfering with regular survey operations. There was no evidence of a mode effect, i.e. no differences between online or paper, and this has meant support for a wider roll out of the electronic questionnaire to other business surveys. All of this work is under the radar, questionnaire and survey design but essential for providing the foundation for the GSS strategic objective to collect more data electronically.   
     
   - We have also redesigned the sample for the International Passenger Survey in order to improve the precision of national migration statistics whilst maintaining that of expenditure estimates – also produced from the IPS. Survey methodologists have ensured that quality standards have been maintained for this high profile output through optimally allocating a new sample design.   
     
   - ONS Methodology have collaborated with leading academics to develop small area estimation approaches that combine survey, Census and admin data in a modeling and statistical framework. This allows us to produce estimates of variables at more detailed levels than previously possible from just survey data. For example we have produced estimates of mean income at the MSOA level using a regression approach whereas direct survey estimates can only be reliably produced at the regional level. And then responding to a direct user and policy need we developed small area estimates of households in poverty at the MSOA level – since the policy need is more about the distribution of rather than average income. We’ve also developed small area estimation structural models to produce population estimates broken down by ethnic group at the LA level – before this was only possible every 10 years using Census data. An example where we are sharing this work at an early stage as a step towards implementation. These approaches will be critical moving forward to meet the demand from users for more detailed statistics and as we make more use of alternative data sources (such as admin and big data) – the real key to making the best use of these data are combining them with survey and Census data to produce reliable estimates in a statistical sound framework – which allows us to understand, measure and communicate the quality of these estimates through the derivation of confidence intervals.   
     
   - The Census is one of ONS’ most high profile activities that takes place every 10 years. A lot of focus is necessarily given to the publically facing data collection exercise and also the outputs when they are published but there is a huge amount of methodological expertise that goes into producing the outputs for users; from designing the questionnaire, supporting the development of the address register the infrastructure that underpins the collection exercise, to edit and imputation, coverage survey design and subsequent estimation and adjustment, design of the statistical geographies and classifications used to design outputs to the statistical disclosure control techniques that will ensure the confidentiality of the estimates released. Often these key methodological steps go unnoticed until there are issues and hence increased scrutiny of the estimates. However, many of the methodologists working on the Census are international experts, for example the methods proposed for SDC for Census have recently been adopted by Eurostat and other international statistical offices – we are leading the way. Going back to my point about gold plated/ivory towers disclosure control is definitely an area where we wouldn’t want to cut corners – particularly since if we get it wrong John Pulllinger, National Statistician could go to prison – don’t like to think about what the equivalent would have been in Roman times.  
     
   - So a huge amount of methodological work underpins the Census. And this is just for the 2021 Census. The ONS are also investigating the feasibility of undertaking an Admin Data Census in the future which will require new methodological approaches to combining, linking and estimating from administrative and survey data. Huge statistical challenges that will require methodologists to address.  
     
   - A key challenge for the Admin Data Census and more generally as we make more use of alternative data sources is the challenge of measuring the uncertainty of estimates produced using these new data sources. We have standard approaches and statistical theory for measuring the uncertainty of estimates produced from surveys. But what about estimates produced from admin or big data sources where we might have limited knowledge about the quality of the data sources? How do we meet our obligations within the Code of Practice in order to communicate whether estimates are fit for purpose. Within ONS Methodology we have been working with academics, working collaboratively we’ve been investigating simulation approaches to produce measures of uncertainty for population estimates – which are produced from registers, admin, survey and Census data. A confidence interval for an estimate may not be the most exciting or interesting statistic (perhaps again like sanitation) – but it provides essential information about the quality of the statistic and allows users to assess whether they are fit for purpose and to identify statistically significant differences between areas and across time. We are the first NSI to produce measures of uncertainty for population estimates.  
     
   - As well as investigating the use of admin data within population and social statistics methodologists have also been integral to incorporating admin data within economic statistics – we have working in a cross disciplinary team involving IT and business area specialists integrating VAT into National Accounts – developing new methods to process and integrate this new data source whilst ensuring the quality of outputs.  
      
   - ONS Methodology have also been investing in research around big data and data science techniques and how this might be applied to improve the accuracy, timeliness, relevance of official statistics. Many examples. Just one being the use of natural language processing techniques to improve the efficiency of analysing text when processing data to produce outputs – for example currently we have to process large amounts of text responses to the British Crime Survey – investigating how NLP can create efficiencies by automating some of these manual processes. This work has the potential to improve processing times but also the quality of outputs –but is hidden ‘under the bonnet’. In addition the Big Data team have led the way in developing web scraping approaches to data collection but also developing methods to process and classify the data using machine learning approaches – leading to the production of experimental daily (rather than monthly) CPI based on these data. So an example where we have released research or experimental outputs early rather than waiting for final production. And we are supporting this work moving more into implementation. In addition we have developed guidance for web scraping – ensuring approaches are ethically sound and leading best practice across Government.  
     
   - And my final example is one where methodologists, data scientists and IT developers have been working in a multi-disciplinary team to develop an address and matching service for use across ONS and wider across Government. Underlying this service are approaches in matching between addresses across different data sources that make use of sophisticated methodological and data science techniques, tools and technologies. All ‘under the bonnet’ of the flashy tool/service that gets demonstrated.
9. Hopefully just this handful of examples demonstrates the range of different innovative and important work that methodologists are undertaking to support and improve the production of official statistics. Just like the Romans and their roads, sanitation, medicine and public baths! Also hope that the examples demonstrate that methodologists are experts in their field with international reputations that are taking on the really challenging technical problems that perhaps others avoid or gloss over. Note – taken from ONS Methodology but I know there will be lots of other examples from across the GSS – many of which will be presented today.
10. So in summary we need to improve the way we communicate methodological work, we should be more confident about sharing early work in order to get feedback – I’ve highlighted some areas where we’ve done this, at times we need to focus more on delivery and implementation and again more examples where we are working collaboratively with colleagues from other disciplines to do this. Methodologists also need to be pragmatic about what is ‘fit for purpose’ whilst ensuring the ONS reputation for high quality and trust worthy statistics is maintained.  
    Only this way will we achieve the 4 I’s that make up the theme for this conference: Insight, Innovation, Implementation and Impact.  
    Coming back to the Monty Python sketch, I hope I have answered the question around ‘what have the methodologists ever done for official statistics’ through the range of examples of excellent work I’ve presented. Actually what I hope I’ve also shown is that it isn’t and it shouldn’t just be about the methodologists – but the methodologist working as part of a multi-disciplinary team ever done for official statistics. And we’ve done loads – we should we proud of our work and we should shout about it and celebrate it.
11. So use opportunities like the GSS Methodology Symposium and other forums to showcase, discuss and learn more about methodological work that is being undertaken across all of the GSS. I think it shows that it isn’t until we sit back and think and celebrate and shout about our work that we recognise the important and innovative work that is going on and is essential to underpin official statistics, particularly in light of greater use of new data sources and the challenges that brings. We won’t be able to meet this challenge for official statistics without the methodologists.

Jane Naylor

July 2017