



**UK Statistics  
Authority**

**Fifth Meeting of the  
National Statistician's Data Ethics Advisory Committee**

---

**Minute, Agenda and Papers**

---

Tuesday 5 July 2016

10:30 – 15:00

Board Room, UK Statistics Authority  
London



## **National Statistician's Data Ethics Advisory Committee**

### **Minute**

**Tuesday, 5 July 2016**  
**Board Room, Drummond Gate, London**

#### **Present**

##### **Members**

Mr Ian Cope (Chair)  
Mr Robert Bumpstead (Deputy Chair)  
Mr Colin Godbold  
Ms Annie Hitchman  
Ms Isabel Nisbet  
Ms Marion Oswald  
Professor Martin Severs

##### **UK Statistics Authority**

Mr Adil Deedat  
Dr Simon Whitworth

##### **Office for National Statistics**

Mr Jon Wroth-Smith (for item 3 and item 7)  
Mr David Johnson for item 5  
Ms Jen Woolford for item 6  
Mr Peter Stokes for item 8

##### **Cabinet Office**

Mr Oliver Buckley (for item 4)  
Ms Madeleine Greenhalgh (for item 4)

##### **British Broadcasting Corporation**

Mr Mark Chipperfield (for item 7)

##### **Apologies:**

Dr Dean Machin  
Mr Neil McIvor  
Mr Osama Rahman  
Mr Hetan Shah

#### **1. Minutes and matters arising from the previous meeting**

- 1.1 The Chair welcomed members to the fifth meeting of the National Statistician's Data Ethics Advisory Committee (NSDEC).
- 1.2 The Chair informed the meeting that the minute of the fourth meeting had been agreed by correspondence. The minute, agenda and papers from the last meeting are now published on the [UK Statistics Authority website](#).
- 1.3 The Secretariat updated the meeting with progress on actions from the previous meetings. Most actions were complete or in progress and the secretariat updated the committee accordingly.

## 2. Chair's report

- 2.1 The Chair informed the meeting of ongoing engagement to raise the profile of NSDEC and its work to date. Members heard that the Chair had presented at the trust, risk, information and law conference at the University of Winchester in April. The presentation focused on the use of big data techniques and new sources in the production of National Statistics and the opportunities, ethics and governance challenges associated with these.
- 2.2 The Deputy Chair and Mr Deedat had presented to Heads of Profession for Statistics (HOPS) earlier this month. This included informing HOPS of the work of NSDEC and outlining examples of projects which had been considered.
- 2.3 The Chair presented at a recent conference on ethical issues in big data hosted by Cancer Research and Winton Capital. The presentation focussed on ethical issues in data linkage and the measures being taken by the UK Government Statistical Service (GSS) to ensure robust oversight of data linkage proposals. The Chair also used examples of data linkage projects considered by NSDEC to illustrate ethical oversight in practice.
- 2.4 The meeting heard that engagement across government about the work of NSDEC will continue over the summer. The Deputy Chair will be presenting to the GSS Statistical Policy and Standards Committee, while the Chair will be presenting to the Departmental Directors of Analysis.
- 2.5 The Chair provided the meeting with an update on projects previously considered by the committee. At the last meeting the Office for National Statistics presented a proposal to use paradata to select participants for follow up research to understand experiences of respondents. Members agreed that the proposal could proceed subject to a consent form and information sheet being provided to the committee for consideration. Since the last meeting members have commented on both of these via correspondence, and have approved subject to comments being implemented.
- 2.6 The Chair informed the meeting that revisions recommended to Welsh Government in relation to their proposal to share data with the University of Bath have been implemented. Members heard that the fair process notice and associated regulations governing the collection of data allow for the sharing of the information requested from the University of Bath. As far as privacy is concerned, the Welsh Government will also be anonymising the school details relating to pupil records and will now be providing the data at middle super output area rather than lower super output area. In addition information at Key Stage 4 will now be provided as aggregate indicators of attainment rather than the results of each individual subject. This will further minimise the risk of an individual being identified. Welsh Government also confirmed that the research proposal itself has now been deliberated on by the University of Bath's Ethics committee, with written confirmation approving the research now received by Welsh Government.
- 2.7 The meeting heard that the Scottish Government have begun disseminating their methodology for linking data. Engagement is taking place with partners in the Administrative Data Research Centre in Scotland. Scottish Government will also share their methods with the Economic and Social Research Council and other centres in the Administrative Data Research Network.
- 2.8 Members heard that the recommendations made to an initial proposal relating a study of end of life care for children had been fed back to the University of Southampton. The University of Southampton are in the process of bidding for funding to undertake the work. Once this has been secured an application will be submitted with members initial recommendations implemented.

- 2.9 The Chair updated members on progress with identifying a suitable access arrangement for the ethnicity from names tool. A range of software has been identified which would lock the tool to an individual's computer preventing onwards transfer and would be time stamped leaving the tool redundant after the allocated time has elapsed. Potential users of the tool would need to demonstrate a clear public benefit from its use before access is granted. Once these arrangements have been finalised and a memorandum of understanding has been agreed between UCL and ONS, these will be circulated to members for further consideration.

**Action: Secretariat to circulate memorandum of understanding to Committee when available.**

### **3. Data Sharing**

- 3.1 The Chair invited Mr Jon Wroth-Smith to present this item. Mr Wroth-Smith informed members that, following the last meeting, he had implemented comments made by the Committee relating to how the data shares were presented.
- 3.2 The Committee were satisfied with the changes made to the data shares annex and suggested that greater information relating to the potential uses be included in future updates.
- 3.3 Members considered whether NSDEC would in the future deliberate on proposals to access particular data sets. While most data being acquired by the office may not be contentious in nature, some data sets such as those relating to health or income may need additional consideration.
- 3.4 The Chair clarified the role of NSDEC with regards to data shares. The meeting agreed that access proposals relating to health or income data should be considered. It was suggested that the acquisition of such data be subject to full ethical review at a future meeting.
- 3.5 As ONS considers exploring acquisition of commercial data, NSDEC suggested that these data also be included within the annex at future meetings.

**Action: Secretariat to work with the Administrative Data Division to bring an application relating to acquisition of income data**

### **4. Public views on government use of data science**

- 4.1 The Chair welcomed Mr Oliver Buckley and Ms Madeleine Greenhalgh from the Government Digital Service (GDS) to the meeting.
- 4.2 Mr Buckley introduced the work his team has been doing to improve data quality, literacy and infrastructure across government.
- 4.3 Ms Greenhalgh then provided an overview of public engagement work undertaken by IPSOS Mori to understand public views on use of data science by government. The research was led by the Government Digital Service and part funded by Science-wise and the Office for National Statistics. It included four sets of focus groups in different locations across the country and online surveys.
- 4.4 The meeting heard that case studies were used to facilitate public understanding of data science. It was reported that there was generally a broad acceptance of the use of data science by government once risks and benefits have been explained.

- 4.5 The presentation fed back public views on a hypothetical case study from ONS, which set out how mobile phone data could be used in the production of population estimates. The meeting heard that there was general acceptance that ONS use of such data was relatively uncontentious. Participants understood the public benefit of population estimates to government and most saw the advantage of using location data, once the method had been explained.
- 4.6 Some participants expressed concerns with using mobile phone location data in isolation and questioned whether modelling techniques were sophisticated enough to produce accurate statistics. The biggest concern however was around the delivery of future public services, which are informed by ONS statistics and the negative impact on individuals of using inaccurate data to change public services.
- 4.7 Mr Buckley then provided the committee with an overview of ongoing work, communicating with the public to establish a contract between the citizen and the state for the use of public data.
- 4.8 It was suggested that there will be breaches of data and that any future framework developed by GDS needs to explicitly recognise this.
- 4.9 Members suggested that NSDEC consider how the ethical principles within the GDS framework fit with NSDEC's own principles.

**Action: Secretariat to compare and contrast NSDEC ethical principles with GDS ethical principles and report back to NSDEC.**

## 5. Centre of excellence for data science

- 5.1 Mr David Johnson from Digital Services Technology and Methodology Directorate in ONS outlined the office's plan for a data science campus. Mr Johnson informed the committee that the data infrastructure and recruitment for the campus were underway and would be finished later in the summer. Given the short timeline to delivery Mr Johnson appreciated the need to balance the work of the data science campus with existing ethical structures.
- 5.2 Mr Johnson informed the Committee of relevant recommendations from the Bean review around the use and capability of data science and the need for ONS to increase its skills, understanding and application of new methods and data in producing statistics.
- 5.3 The meeting heard that the campus will be partnered with a number of academic institutions including the Universities of Bath, Cardiff and Lancaster. ONS and the GSS will work with these partners to deliver an MSc programme and PhDs. By 2017 it is expected that the campus will have between 30 and 36 staff which will rise to 100 over a five year period. It will also look to re-skill existing ONS staff and offer apprenticeships to school leavers.
- 5.4 The campus will consider how other data could be incorporated in to the production of official statistics. For example data from Uber and Airbnb could be used in the production of economic statistics.
- 5.5 Early projects within the campus are likely to include research on educational attainment and its impact on longevity. This would make use of historic education data, graduate surveys and ONS data such as the Census.
- 5.6 Mr Johnson concluded by seeking advice from NSDEC on how the volume of data science projects could be accommodated by existing governance structures.
- 5.7 Members thanked Mr Johnson for a comprehensive presentation. It was suggested that information relating to those projects that are likely to begin before the next meeting will be circulated to members for consideration by correspondence.

- 5.8 The next meeting will be held in Newport, and members will be able to visit the data science campus and get more detailed perspective of the work.

**Action: Secretariat to provide initial projects for the data science campus by correspondence should they begin before the next meeting.**

**6. Draft safeguarding policy and guidance [NSDEC(16)09]**

- 6.1 Ms Jen Woolford from the Life Events division in ONS introduced the paper. Since the last meeting NSDEC comments had been implemented and the draft policy and guidance have been discussed at the GSS Statistical Policy and Standards Committee and at a meeting of the National Statistics Executive Group. It was reported that the policy and guidance would be piloted later this year.
- 6.2 Members welcomed the changes made to the draft policy and guidance. It was suggested that the policy and or guidance should include information on how staff would be supported following initial reporting of a safeguarding incident. Training should also be provided to staff.
- 6.3 Members suggested that incidents be rehearsed at different times, for example during out of office hours.
- 6.4 Members invited Ms Woolford to return a future meeting to update the committee on the pilot.

**7. Application: Television licensing prosecutions: gender disparity [NSDEC(16)10]**

- 7.1 The Chair welcomed Mr Jon Wroth-Smith from the Administrative Data Division in ONS and Mr Mark Chipperfield from the British Broadcasting Corporation to the meeting.
- 7.2 Mr Wroth-Smith introduced the item and explained that following an independent review of TV licensing a gender disparity had been identified, where considerably more women compared with men were prosecuted for not having a TV license.
- 7.3 Mr Chipperfield informed the meeting that the results of the analysis would provide more information on the groups impacted. This would help the BBC to better understand reasons for the disparity and consider whether current stakeholder activities are sufficient and whether training materials require updating.
- 7.4 Members sought clarity around consent and suggested the BBC further explore data sharing in line with their privacy statement. The meeting heard that only aggregate anonymised data would be shared with the BBC.
- 7.5 It was suggested that the BBC may wish to review the proposed contract with ONS and the BBC's third party. As it currently stands ONS would be a data processor; however members suggested that the BBC may want to show ONS as the controller.
- 7.6 Members agreed that the proposal could proceed subject to minor revisions.

**Action: Secretariat to work with Administrative Data Division and British Broadcasting Corporation to revise the application.**

**8. Changes to the Approved Researcher**

- 8.1 Mr Peter Stokes from the ONS Public Policy Division introduced the item. The meeting heard that following consultation the Approved Researcher process had

been changed. The changes will see researchers trained and accredited once while projects would still be considered on a case by case basis.

- 8.2 The changes to the Approved Researcher process would also see commercial organisations eligible to access ONS data in the Virtual Microdata Laboratory for the first time. Such organisations will however need to demonstrate a clear public benefit from their proposal to use the data.
- 8.3 It was suggested that NSDEC consider proposals for access from commercial bodies to ensure that they are ethical and demonstrate public benefit.
- 8.4 It was agreed that following a number of commercial projects being considered by the Committee precedents could be set which would facilitate deliberation on other proposals similar in nature.

**Action: Secretariat to circulate link to Approved Researcher when published later this month researcher.**

## **9. NSDEC self-assessment**

- 9.1 The Secretariat introduced the item and provided a summary of the outcome of the self assessment.
- 9.2 It was reported that the views of members were generally positive. There were suggestions for the inclusion on the Committee of an individual with a relevant social research background.
- 9.3 It was agreed that between meetings, the secretariat will provide members with an update. This will include:
  - i. relevant media coverage relating to ONS and the UK Statistics Authority;
  - ii. relevant publications;
  - iii. an update on relevant meetings and conferences attended by members; and
  - iv. an update on progress for projects implementing revisions.

**Action: Secretariat to provide updates between meetings**

## **10. Any other business**

- 10.1 There was no other business.



**UK STATISTICS AUTHORITY**

**NATIONAL STATISTICIAN'S DATA ETHICS ADVISORY COMMITTEE**

**Agenda**

**Tuesday 5 July  
Board Room Drummond Gate  
10:30am – 3:00pm**

**Chair:** Mr Ian Cope  
**Apologies:** Dr Dean Machin  
Mr Neil McIvor  
Mr Osama Rahman  
Mr Hetan Shah

**(10:30am to 12:00pm)**

1 10:30am	Minute and matters arising from the previous meeting	Mr Ian Cope
2 10:40am	Chair's report	Oral Report Mr Ian Cope
3 11:00am	Data shares Annex A: acquisition of new data sources	Oral Report Mr Jon Wroth-Smith
4 11.15am	Public views on government use of data science	Oral Report Mr Oliver Buckley Ms Madeleine Greenhaigh

**Lunch (12:00pm to 12:30pm)**

**(12:30pm to 3:00pm)**

5 12:30pm	Centre of excellence for data science	Oral Report Mr David Johnson
6 1:00pm	Draft safeguarding policy	NSDEC(16)09 Ms Jen Woolford
7 1:30pm	Television licensing prosecutions: gender disparity	NSDEC(16)10 Mr Jon Wroth-Smith Mr Mark Chipperfield
8 2:00pm	Changes to the approved researcher	Oral Report Mr Pete Stokes
9 2:30pm	NSDEC self-assessment	Discussion Mr Ian Cope
10 3:00pm	Any other business	

**Next meeting: Tuesday 11 October, 2016**

**Location: Newport**



**National Statistician's Data Ethics Advisory Committee****Minute**

**Tuesday, 19 April 2016  
1103-1105, Titchfield, Fareham**

**Present****Members**

Mr Ian Cope (Chair)  
Mr Robert Bumpstead (Deputy Chair)  
Ms Annie Hitchman  
Dr Dean Machin  
Ms Isabel Nisbet  
Mr Osama Rahman  
Professor Martin Severs  
Mr Hetan Shah

**UK Statistics Authority**

Mr Adil Deedat  
Dr Simon Whitworth

**Office for National Statistics**

Ms Carolyn Watson (for item 3)  
Ms Lucy Vickers (for item 3)  
Mr Andy Teague (for item 4)  
Mr Steve Woodland (for item 5)  
Ms Jen Woolford (for item 5)  
Mr Pete Brodie (for item 6)  
Ms Emma Timm (for item 6)  
Mr Neil Bannister (for item 9)

**Welsh Government**

Mr Stephen Hughes (for Item 7)

**Scottish Government**

Mr Albert King (for item 8)  
Mr Gerald Donnelly (for item 8)

**Apologies:**

Mr Colin Godbold  
Mr Neil McIvor  
Ms Marion Oswald

**1. Minutes and matters arising from the previous meeting**

- 1.1 The minute from the last meeting was agreed by correspondence, signed off by the Chair and is now published on the [UK Statistics Authority Website](#).



Chair's report  
Mr Ian Cope



Data shares

Oral report

Mr Jon Wroth-Smith

## **List of Annexes**

**Annex A   Acquisition of new data sources, Jon Wroth-Smith,  
28 June 2016**





**Annex A – Acquisition of New Data Sources****PAYE and Benefits Information**

Data being sought	Level of Detail	Legal Gateway	Public Benefit/Good of Acquiring
Income data (including pay and benefits).	Record level	No longer being pursued as an Information Sharing Order (ISO). Opportunity to provide a statistical service to HMRC (s22 SRSA).	Has the potential to improve or fill gaps on existing ONS social and business surveys (including the Census). Providing a statistical service would enable ONS to test systems, understand potential quality issues and develop processing ahead of access through generic legislation.

**Valuation Office Agency (VOA)**

Data being sought	Level of Detail	Legal Gateway	Public Benefit/Good of Acquiring
Address information (includes date information for records, number of bathrooms/bedrooms, age of property) to be supplied in July 2016 focussed on Census and ONS social surveys. Will supplement enduring supply established in November 2015 already benefiting accuracy of House Price Index (HPI).	Record level address	Information Sharing Order already established.	Potential to improve the accuracy and content of the address register to be used to underpin the 2021 Census. Potential to reduce burden on census respondents - replacing number of rooms question. Benefits transferable to other ONS social surveys. VOA data recently used to directly improve House Price Index (HPI).

**Corporation Tax**

Data being sought	Level of Detail	Legal Gateway	Public Benefit/Good of Acquiring
Data on corporations (includes turnover, profit, liabilities, employment, pensions).	Record level	No longer being pursued as an Information Sharing Order (ISO). Alternatives being assessed.	Potential to improve measurement of Gross Domestic Product (GDP) and National Accounts through business surveys. Scope to improve sampling frames (notably the Inter Departmental Business Register), sampling methods, quality assurance and as alternative measures.

**National Border Target Centre**

Data being sought	Level of Detail	Legal Gateway	Public Benefit/Good of Acquiring
Data on airline carrier, overseas origin airport, crew.	Aggregate	Common Law.	Potential to improve sampling frame and weighting used in the International Passenger Survey (IPS). Also being explored for use in migration/population statistics (beyond the IPS).

**Health Demographics (Hospital Episode Statistics (HES))**

Data being sought	Level of Detail	Legal Gateway	Public Benefit/Good of Acquiring
Demographic data on date of birth, sex, ethnicity plus activity information (appointment, admission, discharge dates).	Record level	The acquisition process for this is through Data Access Request Service (DARS).	Potential for identification of 'signs of life' for administrative data based population statistics as part of Census Transformation Programme and for enumeration of the 2021 Census itself. Improvements in matching algorithms (where NHS number is missing). Improvements to accuracy/granularity of mortality analysis.

**Financial Conduct Authority (FCA)**

Data being sought	Level of Detail	Legal Gateway	Public Benefit/Good of Acquiring
Data on Enhanced Financial Accounts, Regional Accounts, Financial Enquiries, Public corporations.	Record Level	No longer bring pursued as an Information Sharing Order (ISO). Alternatives being assessed.	Improve the quality of the National Accounts through the Enhanced Financial Accounts project.

**Solvency II**

Data being sought	Level of Detail	Legal Gateway	Public Benefit/Good of Acquiring
Enhanced financial accounts covering flows, transactions and stocks for a detailed breakdown of institutional sectors and financial instruments.	Record Level	No longer being pursued as an Information Sharing Order (ISO). Alternatives being assessed.	The development, production and dissemination of the UK's financial accounts, to improve the quality of the financial accounts for the insurance corporations sector.

Public views on government use of data science

Oral report

Mr Oliver Buckley and Ms Madeleine Greenhalgh



Centre of excellence for data science

Oral report

Mr David Johnson



**UK Statistics Authority**  
**National Statistician's Data Ethics Advisory Committee**  
NSDEC(16)09  
***Draft Safeguarding Policy***

This paper is in draft and will be published in due course





## UK Statistics Authority

### National Statistician's Data Ethics Advisory Committee

NSDEC(16)10

#### *Television licensing prosecutions: gender disparity*

#### Purpose

1. This paper presents a proposal from the British Broadcasting Corporation (BBC), who has requested the Office for National Statistics (ONS) to undertake analysis on its behalf using the 2011 Census. The aim of this research is to inform the BBC of whether a gender disparity in prosecutions, identified in a recent review, is the result of women being disproportionately targeted.

#### Recommendations

2. Members of NSDEC are invited to consider the project application at **Annex A** and advise the National Statistician to:
  - i. approve the proposal and allow it to proceed;
  - ii. approve the proposal subject to minor revisions;
  - iii. recommend major revisions to the proposal and request the proposal be resubmitted to a future meeting once implemented; or
  - iv. reject the proposal advising that the proposal be stopped from proceeding.

#### Background

3. The Department for Culture Media and Sport recently commissioned an [independent review](#)<sup>1</sup> of current sanctions in place for failure to hold a TV licence.
4. The review, led by David Perry QC, was published in July 2015 and explored whether there is any inbuilt bias in the process of investigation and prosecution, which unfairly impacts on women or other groups, such as those on low incomes.
5. It was reported that statistics compiled by the Ministry of Justice showed more women than men are prosecuted for TV licence offences. BBC analysis of TV licensing data confirms this with more women than men visited during the enforcement cycle and subsequently prosecuted at the end of the cycle.
6. In order to further explore and understand the reasons for this gender disparity in prosecutions, the BBC has requested ONS undertake some analysis to help further understand the reasons for this trend.
7. The proposal would see the BBC provide ONS with TVL address lists relating to:
  - i. addresses eligible for a visit;
  - ii. names and addresses where a statement has been taken; and
  - iii. a random ten per cent sample of names and addresses from the TV License holder database.
8. ONS would then link these addresses to the 2011 Census to retrieve information relating to household structure and economic activity. ONS would aggregate this information to a non-disclosive level of geography to ensure privacy is protected before passing the anonymised aggregated data to the BBC. The names information provided on two of the TVL lists will be used to assess whether the household structure is the same as it was in 2011.

<sup>1</sup>[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/445212/166926\\_Perry\\_Review\\_Text-L-PB.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/445212/166926_Perry_Review_Text-L-PB.pdf)

9. The findings of the analysis will further clarify the disparity in prosecutions between men and women, identified within the Perry review. The BBC will use the analysis to consider how the disparity may be addressed. This will benefit the public by ensuring that BBC, a public body, prosecutes equally.

**Adil Deedat, Data Governance and Policy, UK Statistics Authority, 23 June 2016**

#### **List of Annexes**

**Annex A      Application: Television licensing prosecutions: gender disparity, Mr Stephen Milner, Office For National Statistics and Mr Mark Chipperfield British Broadcasting Corporation**



## National Statistician's Data Ethics Advisory Committee

### Application for Ethical Review

#### The Application Process

This is an application form for applying for ethical review from the National Statistician's Data Ethics Advisory Committee (NSDEC). You should use the additional guidance when completing this form.

The application form should be completed in **plain English** which is understandable to lay members and all abbreviations should be explained the first time they are used. The form should contain sufficient information to ensure a thorough ethical review can take place.

Please word process the form using Arial or Times New Roman font, size 11. Where necessary expand text boxes on the form to accommodate answers, but ensure word counts are adhered to where specified.

Where sections are not relevant to your study please mark as N/A.

On completion the responsible owner should sign the application form and send to:  
[nsdec@statistics.gsi.gov.uk](mailto:nsdec@statistics.gsi.gov.uk)

7.1



**Section A**  
**Application Details**

<b>A1</b>	<b>Responsible Owner</b>		
Full Name: [REDACTED]		Position: Head of Administrative Data Delivery	
Address: ONS Titchfield		Email: [REDACTED]	
		Telephone: [REDACTED]	
		Organisation: ONS	
<p><b>Declaration to be signed by the responsible owner</b></p> <p>I have met with and advised the applicant on the ethical aspects of this project design <i>(applicable only if the responsible owner is not the Applicant).</i></p> <p>I understand that it is a requirement for all researchers accessing the data to have undergone relevant training and to have either relevant security clearances or approved researcher status in order to access the data.</p> <p>I am satisfied that the research complies with current professional, departmental and other relevant guidelines.</p> <p>I will ensure that changes in approved research protocols are reported promptly and are not initiated without approval by the National Statistician's Data Ethics Advisory Committee.</p> <p>I will provide notification when the study is complete if it or fails to start or is abandoned.</p> <p>I will ensure that all adverse or unforeseen problems arising from the research are reported in a timely fashion to the National Statistician's Data Ethics Advisory Committee.</p> <p>I will consider all advice received from the National Statistician's Data Ethics Advisory Committee and should I be unable to implement any of the recommendations made, I will provide reasoning in writing to the Committee.</p>			
Print Name: [REDACTED] Signature: [REDACTED] Date: 27 June 2016			



<b>A2 Applicant Details</b> (if applicant is not the responsible owner)	
Full Name: [REDACTED]	Position: Data Integration Lead
Address: ONS Titchfield	Email: [REDACTED]
	Telephone: [REDACTED]
	Organisation: ONS

<b>A3 Project Information</b>	
Project Title: Television licensing prosecutions gender disparity	
Start Date: July 2016	End Date: August 2016
Project Sponsor	
<input checked="" type="checkbox"/> ONS <input type="checkbox"/> ADRN <input type="checkbox"/> GSS <input type="checkbox"/> Collaboration <input checked="" type="checkbox"/> Other British Broadcasting Corporation	

<b>A4 Collaboration and Sponsors</b>	
List of Collaborators/Sponsors	Details and relevant documentation relating to collaboration (you may attach copies of relevant documentation)
British Broadcasting Corporation	The BBC will pay for the provision of data


**A5 Proposed Site of Research** (select all that apply)

**Where will the research take place?**

- ☒ ONS  
☐ VML  
☐ HMRC Data Lab  
☐ ADRC-England  
☐ ADRC-Northern Ireland  
☐ ADRC-Scotland  
☐ ADRC-Wales  
☐ Other (please specify).....

**Is this a secure site?**

- ☒ Yes   ☐ No

## Section B

### Project Details

**B1 Please provide a brief high level summary of the research giving necessary background** (*max 500 words*)

The Department for Culture, Media and Sport commissioned an independent review of the current sanctions in place for failure to hold a TV licence. The Perry review<sup>1</sup> (led by David Perry QC) was published in July 2015 and found that:

“116. In its consideration of the current sanctions regime, the Review explored whether there is any inbuilt bias in the process of investigation and prosecution, which unfairly impacts on women or other groups, such as those on low incomes.”

<sup>1</sup> <https://www.gov.uk/government/news/independent-review-on-tv-licence-enforcement-published>

"117. Statistics compiled by the Ministry of Justice demonstrate that more women than men are prosecuted for TV licence offences. The BBC also provided statistics, which reveal that of those interviewed in the course of the enforcement process 67% are female, and of those prosecuted 70% are female. These figures are stark and clearly indicate that women are significantly more at risk of prosecution than men. "

A series of recommendations were made by the review including:

*The BBC and the Department for Culture, Media and Sport should explore ways to investigate and consider the gender disparity in TV licence prosecutions."*  
(Recommendation 3).

BBC analysis of Television Licensing (TVL) transaction data confirms more females have prosecution statements taken from them on enforcement visits and therefore more females are prosecuted. At present, the BBC does not know why this is the case. This is in contrast to the fact that there is almost no gender disparity amongst licence holders before the enforcement cycle begins.

The Department for Culture Media and Sport responded to the review recommendations in October 2015. On recommendation 3, the response stated that:

*"Government will work with TV Licensing to explore this disparity, and this is something that will be taken into consideration through Charter Review."*<sup>2</sup>

This proposal therefore aims to further explore reasons for gender disparity in TV licensing prosecutions by using TV Licensing address data and 2011 Census data. The work would see ONS receive three different TV Licensing address data files from the BBC (relating to 2014-2015) – two of which will also include name information (see section B2). These files will be linked by address to the 2011 Census by ONS. Relevant socio-economic information would then be pulled from the 2011 Census such as household structure. These data would be anonymised, aggregated and statistically disclosure controlled as defined in the 2011 Census Outputs policy<sup>3</sup>, before being passed to the BBC. The names information will then be used to assess whether the household is still structurally the same as it was in 2011.

Should the make-up of households visited be proven to be disproportionately female, the BBC would need to consider what actions, if any, could be taken. It is currently not known whether demographic or cultural factors are influencing the disparity. For example, are more female-led households financially challenged and therefore unlicensed or are men in unlicensed households less likely to give a statement to a TV Licensing enquiry officer. When research results are known and if actions are within the BBC's control, actions can be considered. Depending on outcomes, the BBC could look further at its communication and training processes within the context of the law being consistently and fairly applied.

ONS legal services have advised that the proposal is consistent with the provision of statistical services, which is detailed under section 22 of the Statistics and Registration Service Act (2007).

<sup>2</sup> [http://www.bbc.co.uk/bbctrust/governance/regulatory\\_framework/charter\\_agreement.html](http://www.bbc.co.uk/bbctrust/governance/regulatory_framework/charter_agreement.html)

<sup>3</sup> <http://www.ons.gov.uk/census/2011census/howourcensusworks/howwetookthe2011census/howweplannedfordatadelivery/protectingconfidentialitywithstatisticaldisclosurecontrol>

<b>B2</b>	<b>Data Use</b>			
Type of data	Data Level			
	Please specify the name of the data set			
	Aggreg ate Data	Identifiable Data	De-identified personal data	Anonymised/ pseudo anonymised
<b>Administrative data</b> (please specify, e.g. Patient Register 2011, School Census 2012 etc, in the relevant options adjacent)				
<b>Big Data</b> (please specify e.g. Twitter data, smart meters and mobile phones, in the relevant options adjacent)				
<b>Survey Data</b> (please specify e.g. LFS, BRES, etc in the relevant options adjacent)				
<b>Census Data</b> (please specify year, e.g. Census 2011 in the relevant options adjacent)		2011 Census data: names, addresses, household structure, employment		<u>Ready for linking</u> 2011 Census data: names, addresses, household structure, employment
<b>Other</b> (please specify e.g. Ordnance Survey Address register in the relevant options adjacent)		<u>Receive</u> 1. TVL addresses eligible for visit Jan 2014 – Oct 2015 (3m records) 2. TVL names and addresses where statement was taken Jan 2014 – Oct 2015 (566k records) 3. Random 10% sample of names and addresses from TVL licensed database (2.5m records)		<u>Ready for linking</u> 1. TVL addresses eligible for visit Jan 2014 – Oct 2015 (3m records) 2. TVL names and addresses where statement was taken Jan 2014 – Oct 2015 (566k records) 3. Random 10% sample of names and addresses from TVL licensed database (2.5m records)





--	--	--	--	--

<b>B3</b>	<b>How will information be kept confidential and data kept secure?</b>
-----------	--

	<i>(max 500 words)</i>
--	------------------------

TVL data will be passed securely to ONS (see section C3).

Both the TVL data and the required variables from the 2011 Census will be processed and stored in the ONS Secure Linkage Facility (SeLF). This is a CESG<sup>4</sup> accredited environment, which requires hashing (a one-way encryption process that is irreversible) of identifiable data such as names. To further protect the data, separation of duty is implemented. This ensures that those who hash the data, and therefore see identifiable information, do not see attribute data relating to individuals.

A further step to ensure data security is that SeLF is physically disconnected and air-gapped from all other ONS infrastructure. This makes unauthorised access to the data impossible.

All outputs produced within the environment are statistically disclosure controlled before being exported. This ensures that only anonymised aggregate data leaves the environment.

7.1

<sup>4</sup> UK Government's National Technical Authority for Information Assurance (CESG)

<b>B4</b>	<b>Please provide details of the research protocol or methodology</b> (e.g. data linkage, web scraping etc) (max 500 words)
-----------	--

The BBC will provide ONS with three TV Licensing (TVL) data files.

1. Addresses eligible for a visit January 2014-October 2015 (3m)
2. Names and addresses where a statement was taken January 2014-October 2015 (566k)
3. A random ten per cent sample of names and addresses from the TVL license holder database (2.5m)

These will be appended in to a single file with a variable included to indicate the type of address (as listed in points 1-3 above).

The TVL and Census addresses will be linked to enable the following information for each address to be retrieved from 2011 Census data:

- i. **Structure of household:** for example whether female adult only; male adult only; mixed male and female – measuring whether female led households are statistically more frequent at the postcode sector level which is influencing the over-representation of households where a female is interviewed, or interviewed and prosecuted
- ii. **Presence of children in household:** (by female adult only, male adult only, mixed adult households) - as above. The hypothesis is that females are more likely to be interviewed/prosecuted because they are more commonly at home looking after children.
- iii. **Economic activity:** of adults. For example, female looking after family/part time employed/retired; female in full time employment - as above, in addition looking at additional responsibilities or role a female may be playing in a household that could result in a female being more likely to be interviewed/prosecuted. Also any sense of impact of income.
- iv. **Employment Group:** of adults. For example, working/not working could have an effect on a person being more often at home and therefore more likely to be interviewed/prosecuted – as iii above
- v. **Looking for work (yes/no)** – as iv above
- vi. **Age group:** of adults - as above. May provide some supporting impact of this additional variable. This will allow for analysis which will assess whether certain age groups are more likely to answer the door than others. For example, people of retirement age may be more likely to answer the door. The specific age groups used will depend on a compromise between utility of the data for BBC and statistical disclosure assessment on the data.

The BBC has requested information to be aggregated to postal sector level, e.g. PO15 5. However, without undertaking the analyses ONS cannot guarantee this level of geography as it could be disclosive. Statistical disclosure control will determine which level of geography will be provided.

Address data from TVL and the 2011 Census will go through a geo-coding process. This uses the Ordnance Survey's AddressBase to assign a Unique Property Reference Number (UPRN) to each address. The UPRN will be used to link addresses from both sources.

Once addresses are linked the Census addresses can be used to pull characteristic and household level Census data to produce household structure information for each address. These will then be aggregated to a level of geography that meets statistical disclosure



control requirements.

Where addresses are linked between Census and TVL, and where name has been provided in the TVL data, this will be used to check against the name at the Census address. A flag will be derived that denotes whether the name on the TVL data matches any of the names in the address in the 2011 Census data. This will help to produce a measure of confidence in the household being the same in 2014/15 as it was in 2011. This may also be useful in showing a different rate of change in occupants between the TV Licence holder sample addresses and those addresses visited for a statement. This may in turn be extended to the addresses eligible for a visit where no name information is currently held.

For example, we might expect that the addresses which hold a TV Licence may have lower turnover of occupancy as the addresses which don't hold a TV Licence.

Any TVL addresses that don't link to a Census address (either because they were unable to be assigned a UPRN or their UPRN was not found on Census) will be provided to BBC as counts at a suitable level of geography as determined by statistical disclosure control.

**7.1**

<b>B5</b>	<b>Please outline the proposed benefits of the project</b> <i>(max 500 words)</i>
-----------	---

### **Public Benefit**

This proposal will benefit understanding of whether the gender disparity identified in the Perry review is significantly influenced by a pre-existing gender disparity in the make-up of unlicensed UK households, and the extent to which this is a factor in households eligible to be visited by TV Licensing. A clearer understanding of whether there is a significantly higher predominance of female-only or female-led households in the 'eligible for a TVL enforcement visit' base, will help inform considerations as to whether the gender disparity in TV Licensing enforcement outcomes could be addressed. This will be of benefit to the public by ensuring that a public body, i.e. the BBC, can consider additional steps within its control to support women to become correctly licensed and avoid prosecution.

The variables retrieved from 2011 Census, separately and/or in combination, will enable the BBC as far as possible, to establish influencing factors with regard to why there is a disparity in the proportions of the prosecution statements taken by TV licensing, and those eventually prosecuted. Whilst the BBC has examined other data sets available, it believes there is no other more comprehensive publicly available dataset accessible to them that may allow it to specifically understand household profile from a gender perspective, at a level which may assist them in understanding why the gender disparity in prosecutions occurs.

The BBC is pursuing work alongside the proposed work with ONS (including field research with both its own enforcement officers and customers visited) but is not yet aware of any factors within its control that may be impacting this disparity. It believes a significant contributory factor may be the composition of those households. The BBC does not intend to make the findings public but it is nonetheless, clearly in the public interest to investigate this disparity, as recommended by the Perry Review.

As previously noted, the initial requirement is to establish whether the profile of the households visited is one of the primary influencing factors. Once research results are known, the BBC will consider whether there are actions within its control which may mitigate the situation. Those actions could include further communication initiatives via TV Licensing's extensive stakeholder programme for women at a higher risk of prosecution, within the context of fairly and consistently applying the law.

### **Benefit to statistics and research**

The UK's statistical strategy, Better Statistics Better Decisions, focuses on using pre existing data sources to produce survey and Census type statistics. ONS has already had discussions with TV licensing in the past to consider acquiring address related data to understand its statistical quality and usability in the production of population estimates. Due to the nature of the proposal ONS would gain insight into the statistical quality of TV licensing address data which would inform ONS of the usefulness or otherwise of these data for future purposes.

<b>B6</b>	<p><b>Please outline the ethical issues that might arise from the proposed study and how they will be addressed (<i>all research projects have some ethical considerations, so this section must not be left blank</i>)</b></p>
i.	<p><b><i>The use of data has clear benefits for users and serves the public good.</i></b>  This research serves the public good by aiding the BBC in investigating potential gender disparities in TV Licence prosecutions, as advised by independent Perry review (see section B5). Should a disparity be identified the BBC will consider its communications with stakeholders and customers and will explore whether further training of its enforcement team could help to address any disparity found. This would be focussed towards female led households.</p>
ii.	<p><b><i>The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is appropriately considered.</i></b>  All data are protected by hashing (a one way encryption process) identifiers. The work will take place within a CESG<sup>5</sup> accredited environment. Only TV Licensing addresses (and names, where possible) will be sent to ONS. All outputs will be aggregated and statistically disclosure controlled to ensure data passed to the BBC is anonymous. Section 5.2 of the BBC's privacy policy covers sharing personal data<sup>6</sup>. Consent is not required as the BBC is permitted by law to use the information for TV Licensing purposes. This includes contracting third parties to provide information which requires them sometimes to receive customers' data to carry out those contracted services only. Data protection obligations are incorporated in those contracts and the data is not sold or given to a third party other than to provide a service to TVL. ONS Legal Services have advised that all data to be supplied by the BBC are personal.</p>
iii.	<p><b><i>The risks and limits of new technologies are considered and there is sufficient human oversight so that methods employed are consistent with recognised standards of integrity and quality.</i></b>  The environment within which the work will be undertaken is "air-gapped" meaning that is separate from all other ONS systems and does not have an internet connection. The data will be handled by experienced ONS researchers (all with a minimum of Security Clearance) in a secure environment, with protective procedures adhered to. There is human oversight at each stage and segregation of duty ensures no one person can view all the unhashed data in one setting.</p>
iv.	<p><b><i>Data used and methods employed are consistent with legal requirements such as the Data Protection Act, the Human Rights Act, the Statistics and Registration Service Act and the common law duty of confidence.</i></b>  ONS Legal services have confirmed this is Legal under section 22 (Statistical Services) of SRSA.</p>
v.	<p><b><i>The views of the public are considered in light of the data used and the perceived benefits of the research.</i></b>  Whilst the views of the public have not been sought with regards to the use of the data in this piece of research the public were consulted as part of the Perry</p>

<sup>5</sup> UK Government's National Technical Authority for Information Assurance (CESG)

<sup>6</sup> <http://www.tvlicensing.co.uk/privacy-security-policies>

review.

- vi. ***The access, use and sharing of data is transparent, and is communicated clearly and accessibly to the public.***

The BBC's privacy statement informs the public, how and for what purposes the public's data may be used.

ONS produce statistics and undertake research which is conducive to the public good as outlined in the Statistics and Registration Service Act. In line with the Code of Practice ONS will make its analysis accessible to all.

In addition the BBC has previously commissioned ONS to produce anonymised aggregate Census tables. As with all commissioned tables these were published<sup>7</sup>.

<b>B7</b>	<b>How will the findings of the research be disseminated?</b>
-----------	---

Tabulated outputs will be produced subject to disclosure control rules. They will include aggregated counts of addresses cross-tabulated with the different attributes of interest and by a level of geography that means the data are not disclosive.

The findings of the research will be communicated within the BBC and to the Department for Culture, Media, and Sport (DCMS) on the issue of gender disparity of TV Licence prosecutions. This would form part of a wider response to recommendations from the Perry Review. There is no current intention to publish the findings of this research nor is it understood the DCMS have current plans to do so. Any proposed publication would be provided to ONS in good time to review anything where the output and/or findings of this work, supported by Census data, were to be referred to.

ONS will make its analysis accessible to all.

<b>B8</b>	<b>Please outline any intended future use for products (such as linked data sets or tools) produced as a result of the research and how they will be accessed.</b>
-----------	--

There is no future intended use for the products by ONS and all data will be destroyed after the retention period. Any learning gained from use of the data may be shared within ONS, e.g. quality of the TVL address data.

<sup>7</sup> [https://www.google.co.uk/?gfe\\_rd=ssl&ei=a8FjV9PRMpKCmKojAP#q=2011+census+ct0296](https://www.google.co.uk/?gfe_rd=ssl&ei=a8FjV9PRMpKCmKojAP#q=2011+census+ct0296)



ONS will agree a retention period with the BBC, typically 4-6 weeks from the provision of the outputs, after which all data will be destroyed.

As previously noted, the BBC will share findings with DCMS but does not have any plans for publication at the present time. If future publication were to be considered, this would not take place without ONS consultation.

**Section C**  
**Details of Data Subjects**

<b>C1</b>	<b>Data subjects to be studied</b>		
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; padding: 5px;">Does the Study include all subsections of the population (i.e. all ages, sex, ethnic groups etc)</td> <td style="width: 50%; padding: 5px;"> <b>No</b>  <i>If no please detail which subsections with justification(s) below</i> </td> </tr> </table>		Does the Study include all subsections of the population (i.e. all ages, sex, ethnic groups etc)	<b>No</b> <i>If no please detail which subsections with justification(s) below</i>
Does the Study include all subsections of the population (i.e. all ages, sex, ethnic groups etc)	<b>No</b> <i>If no please detail which subsections with justification(s) below</i>		
<p>Subsections of the population (including vulnerable groups) the project focuses on:</p> <ol style="list-style-type: none"> <li>1. Addresses eligible for a visit January 2014 – October 2015 (circa 3m)</li> <li>2. Addresses where a statement was taken (and name of person providing statement) January 2014-October 2015 (566k)</li> <li>3. A random 10% sample of TVL licensed database (circa 2.5m)</li> </ol>			
<p>Justification for focusing on these subsections or groups:</p> <ol style="list-style-type: none"> <li>1. These are addresses that the BBC have identified as not having a TV Licence and potentially watching television</li> <li>2. These are addresses that the BBC have visited to take a statement</li> <li>3. These are addresses that have a TV Licence</li> </ol> <p>Focussing on these groups enables a comparison to be made between the household structure of the targeted addresses (1), the addresses that were visited from the targeted addresses (2) and the general population of TV Licence payers.</p>			





<b>C2</b>	<b>Please detail consent given to use data specified in section B2</b>
-----------	--

Under the BBC's privacy policy "consent is not required" <sup>8</sup>

The BBC believes the Privacy Policy covers this purpose as data are being passed to ONS for the purpose of administering the TV licence system. Additionally the BBC believes it has considered this activity carefully and believes it to be appropriate as it seeks to understand in following up on the Perry review recommendations that field activity is fair and proportionate. Other grounds for processing the data include legitimate interests, public sector equality duty obligations and public functions/administering justice.

7.1

<b>C3</b>	<b>If you are using data held by a third party please detail how you will obtain this</b>
-----------	---

ONS complies fully with Government standards for data handling and security. Regardless of the original source of a dataset, data are transferred on CESG-approved encrypted media (media approved by the UK Government's National Technical Authority for Information Assurance (CESG)), with encryption passwords and/or tokens controlled by ONS Security Managers. Under no circumstances are personal data transported without appropriate levels of encryption in place.

The sensitivity of the information in each dataset is identified by the originating organisation and the level of sensitivity is used to determine the level of protection that is assigned to the data when it is transferred. For more information, please see the CESG website: <http://www.cesg.gov.uk/servicecatalogue/CAPS/Pages/CAPS.aspx>

The BBC's primary partners are acting under the Trademark licensing and outsourced nature of TVL. Proximity (the data processing partner doing the analysis work and transferring the data to ONS) are contracting with ONS under the same provisions as the BBC to Proximity contract. ONS will be processing the data on behalf of the BBC and according to its instructions.

<sup>8</sup> <http://www.tvlicensing.co.uk/privacy-security-policies>



## Changes to the approved researcher

Oral report

Mr Pete Stokes



## NSDEC self-assessment

### Discussion

Mr Ian Cope



Any other business