

Seventh Meeting of the National Statistician's Data Ethics Advisory Committee

Minute, Agenda and Papers

Tuesday 24 January 2017

10:30 - 14:00

Board Room, UK Statistics Authority London

National Statistician's Data Ethics Advisory Committee

Minute

Tuesday, 24 January 2017 Board Room, Drummond Gate, London

Present Members

Mr Ian Cope (Chair)
Ms Vanessa Cuthill
Mr Colin Godbold
Ms Annie Hitchman
Mr Neil McIvor
Ms Isabel Nisbet
Ms Marion Oswald
Mr Osama Rahman

UK Statistics Authority

Mr Adil Deedat Dr Simon Whitworth

Office for National Statistics

Mr Peter Stokes (for item 3 to item 6) Mr Matthew Greenaway (for item 7) Ms Lan Benedikt (for item 8) Mr David Johnson (for item 8)

Scottish Government

Dr Andrew Waugh (for item 9)

Apologies:

Mr Robert Bumpstead Mr Keith Dugmore Dr Dean Machin Professor Martin Severs

1. Minutes and matters arising from the previous meeting

- 1.1 The Chair welcomed members to the seventh meeting of the National Statistician's Data Ethics Advisory Committee (NSDEC). The Chair informed the committee that Hetan Shah has stepped down as a member of NSDEC ahead of his appointment ending in July 2017.
- 1.2 Members were informed that the minute of the sixth 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 Chair updated the meeting with progress on actions from previous meetings. Most actions were complete or in progress and would soon be complete.

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 blogged on the ONS internal portal, informing ONS staff how ethical review could benefit business areas, the committee's work to date and feedback he had received from business areas who had previously engaged with the committee.
- 2.2 The Chair provided members with an update on projects previously considered by NSDEC. The meeting heard that the Welsh Government proposal to analyse deidentified linked Census and National Survey for Wales data, had had all minor revisions implemented. The paper has now been published on the UK Statistics Authority Website. ONS and the Secure Anonymised Information Linkage Databank (SAIL) (both Trusted Third Parties within the Administrative Data Research Network (ADRN)) are now working together to link and de-identify the data before it is made available to Welsh Government researchers within the secure environment.
- 2.3 The second proposal from Welsh Government, which aimed to understand the outflow of Welsh speakers from Wales to England by linking Welsh Census Records to the 2011 Census, received major revisions. Members heard that all revisions required from the researcher had been implemented; however ONS are still waiting for consent from the Secretary of State for Health to onwardly disclose de-identified linked Patient Register data to the Administrative Data Research Network. Should the Secretary of State for Health give consent then the application will be reconsidered by NSDEC as it is currently under major revisions.
- 2.4 The Census Transformation Programme (CTP) presented to the last meeting on their work to date producing income estimates. The meeting heard that CTP would be presenting an application to the next meeting detailing their existing work and their future plans to produce household income estimates with greater granularity.
- 2.5 The proposal from the British Broadcasting Corporation (BBC) to link Census and TV Licensing data will no longer take place. Following ethical approval, the BBC decided not to pursue this research.
- 2.6 All recommendations relating to the Ethnicity from names project are almost complete. At the April meeting members will receive a paper from ONS which details how each of the major revisions have been addressed.
- 2.7 The Chair updated the committee on Data Access Legislation which forms part of the Digital Economy Bill. The meeting heard that the Bill will soon be considered in the Lord's Committee Stage.
- 2.8 The Chair concluded his report by providing a comparison of the Government Digital Service's (GDS) ethical framework, principles and scope and processes with those of NSDEC. Members heard that the Chair will continue engaging with those developing a second iteration of the GDS framework as part of his role on the Data Leader's Network.

3. MRP: Formalising PriceWaterhouseCoopers (PWC) request to use ONS data [NSDEC(17)01]

3.1 The Chair invited Mr Peter Stokes to the meeting. This Approved Researcher (AR) proposal had previously been considered by correspondence. Mr Stokes thanked the committee for expediting this application. This was the first AR application from a commercial organisation following the updating of the AR criteria and guidance.

- 3.2 The meeting heard that there had been follow up meetings with PriceWaterhouseCoopers (PWC) and the commissioning body, London First, who are due to publish the report in February.
- 3.3 In light of this application, members suggested that the Approved Research scheme publishes guidance for approved researchers relating to how analyses are communicated as well as publication timetables. Consideration should also be given to how any changes in publication timetables are managed. This should include contingencies that can be quickly put in place when ONS are not notified in advance of changes to publication timelines.
- 3.4 It was suggested that the Microdata Release Panel (MRP) which governs access to de-identified ONS data through the secure Virtual Microdata Laboratory (VML) consider whether the Approved Researcher process should take in to account whether the body from which a commercial researcher originates should also be assessed to determine whether they are a fit and proper organisation.
- 3.5 It was suggested that the MRP also consider asking Approved Researchers to include their quality assurance processes in their proposals.

Action: Mr Peter Stokes to consider including additional guidance for approved researchers as detailed and provide an update at the next meeting.

4. Applications: Policies and process for expedited review [NSDEC(17)02]

- 4.1 The Chair introduced the item. Following expedited review of an Approved Researcher proposal referred by the Microdata Release Panel, the Chair informed the meeting that consideration of some proposals via correspondence may be necessary in order to ensure equality in timeliness of access to ONS data.
- 4.2 Members agreed that option one, where all members are sent the application for review would be most appropriate and worked well in the case of reviewing [NSDEC(17)01].
- 4.3 Members agreed that, in order for any expedited proposal to be formally approved, a sufficient number of members would need to respond in accordance with a quorate meeting.
- 4.4 It was agreed that where expedited review is sought and NSDEC approve the proposal, the application and summary of member's comments will be published as part of NSDEC's next meeting.

5. MRP: Frontier Economics use of ONS data [NSDEC(17)03]

- 5.1 Mr Peter Stokes provided an overview of the Approved Researcher application. The meeting heard that Frontier Economics, on behalf of Tech City UK, would be replicating an analysis undertaken last year which was published within the Tech Nation UK report.
- 5.2 Members suggested that the public good be clarified. They also sought clarity around the methods employed and suggested consideration of other data sources and methods which could provide a more complete picture of the UK Tech Sector. This could include use of Company's House data or employing methods such as text mining.
- 5.3 NSDEC agreed that the proposal could proceed subject to clarifying the public benefit and considering the use of alternative data sources and methods which may provide a more complete picture of the tech sector.

Action: Mr Pete Stokes to amend the application to clarify public benefit and to discuss with researchers the use of alternative data sources and or methods which may provide a more complete picture of the tech sector.

- 6. MRP: Road safety analysis Ltd use of ONS data [NSDEC(17)04]
- 6.1 Mr Peter Stokes introduced the Approved Researcher proposal which would see Road Safety Analysis Ltd (RSA) use ONS mortality data to produce an analysis for Highways England on road related suicides.
- 6.2 Members suggested that the analysis could also benefit from investigating attempted road suicides using other data sources.
- 6.3 NSDEC identified that there was a small risk of harm to researchers using the data who may come across an individual they have known in the data set but were unaware of their cause of death. In order to minimise risk of harm to researchers it was suggested that researchers are made aware of this when they are analysing the data.
- 6.4 Overall members agreed that the proposal demonstrated a clear public benefit. However members were unclear as to why home postcode of the deceased at the time of death was required for this analysis.
- 6.5 NSDEC agreed that the proposal could proceed subject to reviewing the need for home postcode information.

Action: Mr Stokes to clarify the need for home postcode information and share reasoning with the committee via correspondence.

7. Web-scraping guidance [NSDEC(17)05]

- 7.1 The Chair welcomed Mr Matthew Greenaway to the meeting. Mr Greenaway provided an overview of the proposed guidance which had been compiled to formalise good practices in web-scraping for use by ONS staff in the production of statistics and research which serve the public good.
- 7.2 The meeting heard that in order for ONS to undertake web-scraping on a large scale, i.e. scraping small amounts of data from a large number of websites, there were challenges in checking website terms and conditions and gaining consent from website owners.
- 7.3 Members were generally supportive of the guidance, but suggested further work be undertaken on the guidance.
- 7.4 This included considering the use of text mining and use of machine learning techniques in order to automatically check terms and conditions of websites to understand whether they can be legally scraped.
- 7.5 It was suggested that references to personal data needed to be defined with an emphasis that any proposal to scrape personal data from a website be considered by NSDEC before scraping is undertaken.
- 7.6 It was suggested that ONS may want to undertake engagement with relevant businesses whose websites they are most interested in scraping. The guidance should also provide greater clarity as to what is meant by large amounts of data.
- 7.7 It was suggested that greater clarity be sought as to the legal implications of breaching terms and conditions of website owners. Members referenced a case considered in the EU Court of justice. It was suggested that the Big Data team consider the implications of this ruling and related copyright and data protection in redrafting the guidance.
- 7.8 Further consideration should also be given to the types of websites being scraped to ensure they themselves are not unethical and do not pose a risk to the ONS

Action: Secretariat to work with Mr Greenaway to reflect NSDEC's comments in to the application and to bring back to the committee at a future meeting for further consideration.

Action: The Chair in his role as an ONS executive to offer support to establish a task and finish group to further refine the draft guidance.

- 8. Big Data to Tourism Statistics
- 8.1 The Chair welcomed Ms Lan Benedikt and Mr David Johnson from the ONS Data Science Campus to the meeting. The Data Science Campus is set to act as a hub for the whole of Government to gain practical advantage from the wider investment in data science research and help cement the UK's reputation as an international leader in data science.
- 8.2 Ms Benedikt set out existing practices for producing tourism and migration statistics using the International Passenger Survey (IPS). The meeting heard of limitations to the IPS, such as sampling error. There are a limited number of data sources currently available to ONS to supplement estimates produced using the IPS.
- In order to address this, the Data Science Campus is working with other partners in government, academia and from other National Statistics Institutes to consider how a range of open data sources could be used to produce tourism and migration statistics.
- The meeting heard of initial feasibility work using geo-located Flickr data. Here, through the Flickr API, user data is downloaded. All identifiers which could be used to identify an individual are removed and the data is aggregated to show where groups of tourists are localised around particular landmarks. Ms Benedikt stressed to members that photographs uploaded by Flickr users were not downloaded.
- 8.5 Members were divided in their views on the use of geo-located social media data. There were concerns that whilst individuals may have agreed to sharing their location they may not have been informed in providing consent for reuse of social media user's data. Other members also expressed concerns around minors breaching the terms and conditions of use and registering to have accounts.
- 8.6 Ms Benedikt informed members that research using social media data was purely exploratory at this stage and would contribute to further developing methods and understanding how big data could be used in the production Official Statistics.
- 8.7 Members heard that the Data Science Campus would be working with the ONS Big Data team to develop guidelines for how social media should be used by ONS for research and statistics which serve the public good.
- 8.8 NSDEC members noted the difference between initial feasibility research and the use of social media data in the production of official statistics. It was suggested that this research could proceed for the time being; however the Data Science Campus should consider providing a paper or application to NSDEC addressing each of NSDEC's ethical principles with regards to the use of social media data.
- Action: the Data Science Campus and Big Data team to develop guidance for the use of social media data for research and statistics which serve the public good.
- Action: the Data Science Campus to produce a paper or application detailing their plans for use of social media data to produce experimental statistics.
- 9. National Records of Scotland: using public and commercial sector administrative data in the production of population statistics [NSDEC(17)06]

- 9.1 The Chair welcomed Dr Andrew Waugh from the National Records of Scotland to the meeting.
- 9.2 Dr Waugh outlined the NRS's ambition to use data from commercial companies who were listed on Crown Commercial Services; this could include the use of data held by credit reference agencies for the production of population statistics.
- 9.3 Dr Waugh explained that as government administrative data are collected for administrative purposes there were challenges using it in the production of statistics. For instance 30-59 year old males are often over represented on government sources. This may be the result of these groups either not interacting with GP services or who are no longer resident and not notified DWP of a change in their circumstances. Through considering credit reference agency data it may be possible to determine whether or not individuals are still active in the country.
- 9.4 NSDEC saw the benefits such data could bring to the production of administrative data based population estimates; however there were reservations regarding the ethics of commercial companies profiting from the sale of the public's data which had been provided for another specific purpose (credit checking). Consideration was given to whether the consent would only relate to that specific use.
- 9.5 Members suggested further engagement with DWP and HMRC to explore ways in which their data could be used, in a similar way to how credit reference agency data use was proposed. It was also suggested that NRS work closely with ONS and NISRA to ensure a joined up approach, especially around approaches to commercial data sources.
- 9.6 Members agreed that whilst the proposal had clear public benefits more work should be undertaken to explore existing data sources in government. Discussions with other UK statistical institutions should also be undertaken before considering procurement of commercial datasets. NSDEC therefore recommend major revisions to this proposal whilst these conversations take place.

Action: Dr Andrew Waugh to discuss potential of existing government sources and procurement of commercial data sources with ONS

- 10. Any other business
- 10.1 There was no other business.

UK STATISTICS AUTHORITY

NATIONAL STATISTICIAN'S DATA ETHICS ADVISORY COMMITTEE

Agenda

Tuesday 24 January Board Room, Drummond Gate London 10:30am – 2:00pm

Chair: Mr Ian Cope

Apologies: Mr Keith Dugmore

Dr Dean Machin

Professor Martin Severs

(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	MRP: Formalising PWC request to use	NSDEC(17)01
11:00am	ONS data	Mr Ian Cope
4	Applications: Policies and process for	NSDEC(17)02
11:05am	expedited review	Mr Ian Cope
5	MRP: Frontier Economics	NSDEC(17)03
11:20am		Mr Peter Stokes
6	MRP: Road safety analysis Ltd	NSDEC(17)04
11:40am		Mr Peter Stokes

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

(12:30pm to 2:00pm)

7	Web-scraping policy	NSDEC(17)05
12:30pm		Mr Owen Abbott
		Mr Matthew Greenaway
8	Big Data to estimate and explain tourism	Oral Report
1:00pm	and migration statistics	Ms Lan Benedikt
		Mr David Johnson
9	Administrative data based population and	NSDEC(17)06
1:30pm	household estimates	Dr Andrew Waugh
10 2:00pm	Any other business	

Next meeting: Tuesday 25 April 2017 Location: One Drummond Gate, London These minutes were agreed by correspondence and signed off by the Chair and are now published along with the agenda and papers from the previous meeting on the UK Statistics Authority Website

National Statistician's Data Ethics Advisory Committee

Minute

Tuesday, 11 October 2016 Board Room, Drummond Gate, London

Present Members

Mr Ian Cope (Chair)
Mr Robert Bumpstead (Deputy Chair)
Ms Vanessa Cuthill
Mr Keith Dugmore
Mr Colin Godbold
Mr Neil McIvor
Ms Isabel Nisbet
Ms Marion Oswald

UK Statistics Authority

Mr Adil Deedat Dr Simon Whitworth

Office for National Statistics

Ms Meghan Elkin (for item 3) Mr Richard Heys (for item 4) Mr David Johnson (for item 7)

Welsh Government

Mr Martin Parry (for item 5 and 6)

Apologies:

Ms Annie Hitchman Dr Dean Machin Mr Osama Rahman Professor Martin Severs Mr Hetan Shah

1. Minutes and matters arising from the previous meeting

- 1.1 The Chair welcomed members to the sixth meeting of the National Statistician's Data Ethics Advisory Committee (NSDEC). The Chair introduced two new members to the Committee; Vanessa Cuthill, Director of Research and Enterprise at the University of Essex and Keith Dugmore MBE, founder of the Demographics User Group.
- 1.2 The Chair informed members that the minute of the fifth meeting had been agreed by correspondence. The minute, agenda and papers from the last meeting are now published on the <u>UK Statistics Authority website</u>. One amendment from the minute was agreed.

1.3 The Chair updated the meeting with progress on actions from previous meetings. Most actions were complete or in progress and would soon be complete.

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 since the last meeting the Chair had presented to a cross Government meeting of Departmental Directors of Analysis whilst the Deputy Chair presented to the Statistical Policy and Standards Committee. The Secretariat also presented a poster highlighting the service NSDEC offers at the Government Statistical Service Conference.
- 2.2 The meeting heard that NSDEC members Marion Oswald and Hetan Shah participated in a Radio Four programme about "the online identity crisis".
- 2.3 Marion also informed the committee of the recruitment of a PhD student to work on legal, ethical and privacy issues relating to the work of the ONS Big data team. Previous work undertaken by the team has included the use of twitter data to understand flow to work patterns.
- 2.4 The Chair provided the meeting with an update on projects previously considered by the committee. At the last meeting the British Broadcasting Corporation and the Office for National Statistics presented a proposal which would see ONS link Television Licensing data to the Census to further understand gender disparities identified as a part of the a recent independent review of licensing prosecutions. The meeting heard that recommendations made by the committee had now been implemented and the BBC and ONS are in the process of finalising details of the project.
- 2.5 The final draft of the safeguarding policy and guidance for staff has been agreed along with a form to record concerns. Training for interviewers is now being developed and will be delivered later this year in briefing the National Study of Health and Wellbeing: Children and young people, which will take place in January 2017. Interviewers have now been briefed on the safeguarding policy and will apply it across all surveys they work on. Feedback on the policy will be gathered after six months, along with that of the safeguarding contacts and the Chief Safeguarding Officer who will receive and consider concerns. Provision of training for safeguarding contacts and for senior staff who may undertake the role of Chief Safeguarding officer is being explored with the Social Care Institute for Excellence.
- 2.6 The information sheet and consent form for the para-data project has now been agreed. The Survey Methodology and Statistical Computing Division are now looking to implement this within their work.
- 2.7 The meeting heard that the Memorandum of Understanding between University College London and the Office for National Statistics was still being drafted and access arrangements for the tool are being considered. The Memorandum of Understanding and access arrangements will be shared with the committee once they are finalised.
- 2.8 The Chair updated the committee on Data Access Legislation which forms part of the Digital Economy Bill. The meeting heard that the bill was in committee stages with evidence being heard from Professor Sir Charlie Bean and Mr Hetan Shah amongst others.

3. Administrative Data Census Outputs-Income

3.1 The Chair invited Ms Meghan Elkin, who is leading on this work within ONS's Census Transformation Programme, to present this item. The meeting heard that the Census

Transformation Programme was exploring how DWP and HMRC data sources could be used in the production of income estimates.

- 3.2 Ms Elkin outlined the user need for income data and the public benefit such as to better understand deprivation to enable better targeting of resources and services.
- 3.3 The meeting heard that income questions had never been included in the Census as previous Census tests have shown it to cause a reduction in overall response rates. The recent consultation on the 2021 Census showed the public to be more willing for information relating to their income to be derived from administrative data rather than be included in a Census questionnaire.
- 3.4 Ms Elkin outlined existing methods of producing income statistics, the majority of which are derived from surveys. The meeting heard that ONS had access to some employment and benefits data; however there are some types of employment data such as self employment, and some benefits data which ONS have yet to access which could be integral to producing comprehensive income statistics.
- 3.5 The meeting heard that research is being undertaken within the Census Secure Research Environment where all data is anonymised before analysis. To further protect privacy, incomes above an agreed threshold have been aggregated to a common value in order to mitigate against the disclosure of individuals with high incomes.
- 3.6 It was suggested that care would be needed in communicating the initial income estimates produced. Stakeholders should be made aware of any statistical quality issues with the estimates.
- 3.7 Members suggested that as ONS acquire data for 2015 and 2016, consideration should be given to the changes in pension arrangements which could impact on the analysis.
- 3.8 Members were satisfied with the general use of the data for the purpose of producing aggregate de-identified statistics for publication later this year.
- 3.9 The meeting agreed that the Census Transformation Programme should complete an application form detailing their plans to use PAYE, benefits and other data sources to produce granular income statistics. Moving forward, uses of data sources relating to income will be subject to full ethical review.

Action: Census Transformation Programme to develop formal application on use of pay as you earn and benefits data to produce income estimates.

4. Economic Statistics Centre of Excellence (ESCOE)

- 4.1 The Chair invited Mr Richard Heys, Deputy Director within the Office of the Chief Economic Adviser in the Office for National Statistics, to present this item.
- 4.2 The meeting heard that the Bean review had recommended ONS, in conjunction with suitable partners from academia and the user community, establish a new centre of excellence for the analysis of emerging and future issues in measuring the modern economy.
- 4.3 My Heys informed members that ONS will fund ESCOE as an independent research centre, shaped by ONS's priorities for the development of economic statistics. A proposed model for engagement with NSDEC was also outlined which would see ONS engage with the committee when making data available to ESCOE.
- 4.4 The meeting heard that research specification will be agreed in December.

4.5 Members expressed some concerns that ESCOE appeared similar in nature to the Administrative Data Research Network (ADRN). It was suggested that ESCOE engage with relevant parties from the ADRN to learn from their experiences particularly around security accreditation and burdening data owners.

4.6 It was suggested that the successful tender have proposals reviewed by their own ethics committee before consideration by NSDEC.

Action: Richard Heys to meet with Professor David Hand, Chair of the ADRN Board to learn from the work of the ADRN.

- 5. Welsh Government: Exploring differences in the reporting of Welsh Language ability: An exploratory study linking survey and Census data. [NSDEC(16)11]
- 5.1 The Chair declared an interest as information asset owner for the Census and passed chairing responsibilities to his deputy, Mr Bumpstead. Ms Nisbet also declared an interest as part of her role in Qualifications Wales.
- 5.2 Mr Bumpstead recapped NSDEC's role in considering research proposals from government researchers who require access to ADRN infrastructure.
- 5.3 Mr Bumpstead welcomed Mr Martin Parry, a senior Welsh Language officer from the Welsh Government to present this item. Mr Parry informed the committee that as part of the Well-being of Future Generations Act, a key goal was for a thriving Welsh language.
- Mr Parry explained that there were differences between the number of Welsh speakers estimated from the Census and the National Survey of Wales (NSW). Both of which ask identical questions relating to Welsh Language. In order to understand this difference the proposal would look to link the NSW and the Census to understand the demographic groups who report their Welsh language ability differently on each of the sources.
- 5.5 Members suggested that the application should clarify, using non technical language, how confidentiality would be maintained. The length of time for which the data will be retained should also be included and time scales for the research should be revised.
- 5.6 NSDEC agreed that the project could proceed subject to these minor revisions being implemented.

Action: Secretariat to work with Mr Parry to implement recommendations before the proposal proceeds.

- 6. Welsh Government: An Exploratory study for estimating the outflow of Welsh Speakers from Wales to England from the 2011 Census and the Patient Register [NSDEC(16)12]
- 6.1 Mr Parry introduced the item. The meeting heard that the aim of the proposal was to understand the outflow of Welsh speakers from Wales to England. Mr Parry explained that in the context of the Welsh Government trying to create a Wales with a thriving Welsh language, understanding the outflow of Welsh speakers was important.
- 6.2 Mr Parry outlined the proposed method by which the outflow of Welsh speakers could be estimated, by linking the 2011 Census to the Patient Register in England for years following the Census.
- 6.3 Members suggested that the application needed to outline more clearly what the public benefits would be from the proposal and how the findings would be used in evaluating Welsh Government policy. The research protocol should also be clarified.

- 6.4 Members agreed that the application should make clearer, using non technical language, how confidentiality would be maintained. The legal gateway outlined for accessing the patient register should also be clarified to provide assurance that the relevant permissions have been received from the data owner to use the data for the purposes of this study.
- 6.5 The committee therefore recommended major revisions to the application, and sought recommendations to be implemented before a redrafted application is circulated to the committee for reconsideration.

Action: Mr Parry to redraft application, implementing committee recommendations for further consideration.

7. Visit to the Data Science Campus

- 7.1 Mr David Johnson welcomed NSDEC to the data science campus. The meeting heard that the campus was up and running and taking on apprenticeships for a period of two years.
- 7.2 Mr Johnson informed the committee of the initial research themes for the campus which would be completed in collaboration with academic, industry and government partners and which would include
 - i. the modern economy;
 - ii. sustainability;
 - iii. the UK in a Global context;
 - iv. urban future; and
 - v. society.
- 7.2 Mr Johnson reported that he had been in the early stages of discussions with the secretariat about the development of ethical guidelines which could be incorporated into training that is being offered as part of the data science campus work programme. It was suggested that this would be an important piece of work and resource would have to be devoted to it to produce something suitable.
- 7.3 Mr Johnson reported that he has had discussions with various other departments such as the Department for Environment, Food and Rural Affairs, the Department for Culture, Media and Sport and the Department for Work and Pensions who are looking to develop some sort of ethical review process. It was reported that these departments are interested in hearing more about the work of NSDEC. Mr Johnson offered to put the secretariat in touch with relevant contacts in these departments.

Action: Secretariat to continue working with Mr Johnson to develop ethics training for data scientists.

8. Any other business

The meeting heard that member Isabel Nisbet presented at a recent conference of government social research.

Chair's report

Mr Ian Cope

UK Statistics Authority

National Statistician's Data Ethics Advisory Committee

NSDEC(17)01

Microdata Release Panel: Study on the effects of migrant workers on London's economy

Purpose

- 1. This paper presents a proposal for use of ONS data by PriceWaterhouseCoopers (PWC) on behalf of London First¹, to understand the effects of migrant workers on London's economy. This proposal was approved by NSDEC via correspondence and is presented here to formalise member's deliberations via correspondence.
- 2. Members of NSDEC are presented with the project application at **Annex A** for reference purposes. This project was approved via correspondence on 28 November 2016.

Background

- 3. The Microdata Release Panel (MRP) governs access to the secure ONS Virtual Microdata Laboratory, and provides approved researchers access to de-identified ONS data in order to undertake approved projects.
- 4. In 2016, following a public consultation of the Approved Researcher process, it was agreed that where necessary, the MRP would refer applications to use ONS data via the approved researcher gateway to NSDEC for ethical review.
- 5. This project, which aimed to study the effects of migrant workers on London's economy, was the first to be referred to NSDEC by MRP. Due to the short timelines to Article 50 being invoked and the relevance of this proposal to the UK's exit from the European Union there was a pressing need for this work to begin before 2017 and NSDEC's next meeting.
- 6. The project will benefit understanding of effects of migrant workers on the London economy. The research will provide an evidence base to inform government policy and decision making, including the Government's arrangements to invoke Article 50 of the Lisbon Treaty by the end of March 2017 for the UK's exit from the European Union.

Adil Deedat, NSDEC Secretariat, Central Policy Secretariat, UK Statistics Authority, 3 January 2017

List of Annexes

- Annex A: Application: Study on the effects of migrant of migrant workers on the London's economy, Mr Pete Stokes, Microdata Release Panel, Office for National Statistics, 22 November 2016
- Annex B: Summary of discussions and actions via correspondence, Adil Deedat NSDEC Secretariat, Central Policy Secretariat, UK Statistics Authority, 28 November 2016

¹ http://londonfirst.co.uk/



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 <u>plain English</u> 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

Section A Application Details

A1 Responsible Owner	
Full Name:	Position: Head of Data Access and Exploitation
	Email:
	Telephone:
	Organisation: ONS

Declaration to be signed by the responsible owner

I have met with and advised the applicant on the ethical aspects of this project design (applicable only if the responsible owner is not the Applicant).

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.

I am satisfied that the research complies with current professional, departmental and other relevant guidelines.

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.

I will provide notification when the study is complete if it or fails to start or is abandoned.

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.

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.

Print Name:

Signature:

Date 18 November 2016

9. NRS



A2 Applicant Details (if applic	
Full Name	Position Manager – Risk Assurance
Address:	Email:
	Telephone:
	Organisation: Price Waterhouse Coopers LLP

A3 Project Information		
Project Title: Study on the effects of migrant workers on London's economy		
Start Date : 24/11/16	End Date: 22/12/16	
Project Sponsor (select all that apply)		
□ ONS		
ADRN		
□ GSS		
□ Collaboration		
Other (Please specify)London First (http://londonfirst.co.uk/)		

List of Collaborators/Sponsors Details and relevant documentation		
ist of Collaborators/Sporisors	relating to collaboration (you may attach copies of relevant documentation)	
	copies of relevant accumentation)	



A5 Proposed Site of Research (select all that apply)			
Where will the research take place?			
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)			
Price Waterhouse Coopers (PWC) have been commissioned by London First (a not-for-profit consortium of businesses, including a number of important financial institutions http://londonfirst.co.uk/membership/our-members/) to carry out analysis of the effect of migration on the London Labour Market. They will produce an entirely factual analysis (so not political, or opining in the value of migration) and publish this on their website as they have done for other analyses. The Microdata Release Panel (MRP) consider that there are clear public benefits to the work, such as an assessment of the importance of migrant labour to the London economy and their impact and economic contribution.			
The three main scope areas are:			

4

9

NRS



1. Aggregate overseas worker data

- collection, aggregation and analysis of overseas worker data in London
- 2. The importance of migrant Labour to the London economy
- exploring key policy and services planning questions
- 3. Assessment of the overall economic contribution and impact
- such as the Impact of migrants on wages, in different sectors and in a range of roles requiring different skill levels, and on the London housing market.

The research will inform government policy and decision making, including the Government's arrangements to invoke Article 50 of the Lisbon Treaty by the end of March 2017 for the UK's exit from the European Union. The urgency for the research is therefore determined by the imminent timing for triggering Article 50 and the need to present the findings to key policy stakeholders in January and early February 2017 - including the Greater London Authority (GLA), Home Office officials and the Home Office Minister for Immigration - to inform policy and decision making, and service planning for London and other government priorities such as immigration and employment.

The research project was approved by MRP on 16 November 2016 on the grounds that there was a legal gateway to access the data, it was appropriate use of ONS data and a public benefit was demonstrated. As this is the first MRP project being carried out by a commercial organisation, the project is (as agreed) being referred to NSDEC to assess the ethical use of these data by a commercial organisation. If the project were being carried out by a non-commercial organisation, it would have been agreed without NSDEC consideration.



B2 Data Use

Type of data	Data Level Please specify the name of the data set			
	Aggregat e Data	Identifiable Data	De- identified personal data	Anonymised/ pseudo anonymised
Administrative data (please specify, e.g. Patient Register 2011, School Census 2012 etc, in the relevant options adjacent)				
Big Data (please specify e.g. Twitter data, smart meters and mobile phones, in the relevant options adjacent)				
Survey Data (please specify e.g.LFS, BRES, etc in the relevant options adjacent)			Annual Population Survey (APS) all years 2006 - 2016, Annual Business Survey (ABS), all years from 2008 London Business Survey (LBS), 2014	
Census Data (please specify year, e.g. Census 2011 in the relevant options adjacent)				
Other (please specify e.g. Ordinance Survey Address register in the relevant options adjacent)				

B3 How will information be kept confidential and data kept secure? (max 500 words)

Access to data will only take place within the secure Virtual Microdata Laboratory (VML) environment and all outputs will be checked by the VML team prior to release to ensure

Policies



disclosure control and the confidentiality of data subjects is protected. All analysis and use of the data will be within the VML at a secure VML setting at one of the ONS offices. The PWC researchers have been accredited as Approved Researchers and have successfully completed (including an assessment) the training on safe use of data in research environments (SURE). No data or report drafts will be shared with anyone outside of the project team.

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

The LBS/ABS data will be linked and matched at aggregate level with similar characteristics from the APS/LFS within the VML to help identify data for the following categories:

- industry sector
- employment % and type (e.g. Full time)
- salary band
- country of origin
- nationality
- time in UK
- region (e.g. London)
- London borough
- skill of position and role (employment)

This will be used to carry out the three phases of the project:

1.Produce aggregate overseas worker data for

- a) Number in London and in past years, broken down by EU (old and accession countries) and non EU
- b) Number of dependents
- c) Length of time of stay and employment in London
- d) Average length of time spent in London/UK before returning home/becoming residents
- e) Employed/unemployed proportion for EU/ non-EU and individual countries
- f) Employment profile of an EU/ non-EU migrant arriving in London length of time taken to secure work, unemployment periods, single vs multiple jobs.
- g) Employment profile of migrant dependents

2. Evaluate the importance of migrant Labour to the London economy, such as

- a) Overseas workers impact on the London labour market by sector, type of role and skill level by EU and non EU
- b) Sectors reliance on EU/non-EU migrant labour, skilled and unskilled;
- c) how reliance on overseas workers correlates with UK skills/employment/unemployment levels (e.g. if London's IT sector has a high proportion of migrant workers, how does that correlate with those seeking work in the sector)
- d) The contribution of migrant labour to supply chains e.g. construction sector
- e) The impact on the domestic workforce in terms of housing, cost of living (will be supplemented with additional publicly available datasets outside of the VML once the analysis has been completed and outputs checked, e.g. housing prices per borough and

Policies



population growth)

3. Assess the overall economic contribution and impact for

- a) Impact of migrants on wages, in different sectors and in a range of roles requiring different skill levels
- b) Impact on job creation i.e. do migrants create jobs or displace local workers
- c) The costs of benefits to migrant workers
- d) Impact on tax revenues and GDP/GVA overall and per capita
- e) Impact on London housing market

Dependents of migrants will be counted, but not analysed. For instance, PWC are interested that a migrant worker has two dependents, but they will not look at, or analyse, the characteristics of them.

Given that the focus of the research is migrant workers in London, anyone not economically active (of any age), or above working age, will be flagged as such and not included in the analysis.

Any intermediate data outputs will be cleared by ONS and the aggregated data will be combined with other publicly available data to provide further analysis. These may include:

- 1. HMRC tax credits https://www.gov.uk/government/statistics/tax-credit-statistics-on-eea-nationals
- 2. HMRC tax contribution of EEA nationals -

https://www.gov.uk/government/statistics/further-statistics-on-tax-contribution-of-eeanationals-for-2013-to-2014

- 3. DWP National Insurance number allocations to adult overseas nationals entering the UK https://www.gov.uk/government/collections/national-insurance-number-allocations-to-adult-overseas-nationals-entering-the-uk
- 4. NHS workforce by nationality http://www.hscic.gov.uk/pubs/provisionalmonthlyhchsworkforce

The final report will be cleared by ONS to ensure the confidentiality of data subjects is protected. The report will combine statistical summaries and data visualisations in aggregate form, and correlation analysis of data to reach firm conclusions on the data. The final report will be fact-based so that any conclusions are based on the data, rather than assumptions. In accordance with the conditions of access, the report will be published on the London First website within four weeks of its completion and a link sent to ONS.

B5 | Please outline the proposed benefits of the project (max 500 words)

The output will be a clear, factual report which will provide the foundation for developing immigration policy around Brexit.



The project will benefit understanding of effects of migrant workers on the London economy. The research will provide an evidence base to inform government policy and decision making, including the Government's arrangements to invoke Article 50 of the Lisbon Treaty by the end of March 2017 for the UK's exit from the European Union. The urgency for the research is therefore determined by the imminent timing for triggering Article 50 and the need to present the findings to key policy stakeholders to inform policy and decision making, and service planning for London and other government priorities such as immigration and employment.

On 11 January 2017, London First is presenting the results to the Mayor of London, Sadiq Khan. London First have regular meetings with Paul Regan, Home Office Head of Migration and on 17 January 2017 they are discussing the results of the research and next steps with Paul. On 1 February, London First will present the results to the Minister of State for Immigration, Robert Goodwill.

There will be additional public benefits from the research, including: providing an evidence base for decisions which are likely to significantly benefit the UK economy, society or quality of life of people in the UK; to replicate, validate or challenge existing research; and to significantly extend understanding of social or economic trends or events by improving knowledge or challenging widely accepted analyses. The research will contribute to the UK Statistics Authority's Better Statistics, Better Decisions strategy and there may be benefit in developing it into a case study to demonstrate the impact of microdata research.

B6 Please outline the ethical issues that might arise from the proposed study and how they will be addressed (all research projects have some ethical considerations, so this section must not be left blank)

i. The use of data has clear benefits for users and serves the public good

The project will benefit understanding of effects of migrant workers on the London economy. The research will provide an evidence base to inform government policy and decision making, including the Government's arrangements to invoke Article 50 of the Lisbon Treaty by the end of March 2017 for the UK's exit from the European Union. The urgency for the research is therefore determined by the imminent timing for triggering Article 50 and the need to present the findings to key policy stakeholders to inform policy and decision making, and service planning for London and other government priorities such as immigration and employment.

There will be additional public benefits from the research, including: providing an evidence base for decisions which are likely to significantly benefit the UK economy, society or quality of life of people in the UK; to replicate, validate or challenge existing research; and to significantly extend understanding of social or economic trends or events by improving knowledge or challenging widely accepted analyses. The research will contribute to the UK Statistics Authority's Better Statistics, Better Decisions strategy and there may be benefit in developing it into a case study to demonstrate the impact of microdata research.



ii. The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.

The data will be accessed in the secure VML accredited environment and is deidentified to protect the confidentiality of data subjects. The Approved Researcher
scheme will be the legal gateway used to access the data and the researchers have
successfully completed the Safe User of Research data Environments (SURE) training
(and passed the assessment) and will sign and adhere to the Approved Researcher
declaration setting out how they will manage the data and protect the confidentiality of
data subjects in line with the Statistics and Registration Service Act 2007. The
researchers will not be able to take the data outside of the VML and their outputs will be
checked and cleared by the ONS VML branch to prevent disclosure of data subjects.

iii. 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.

London First intend to produce an entirely factual analysis (so not political, or opining in the value of migration). All analysis will be carried out in the VML and all outputs will be checked to ensure the confidentiality of data subjects is protected.

iv. 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.

Access to the potentially disclosive data will be in a secure environment (VML) and via an approved legal gateway (Approved Researcher scheme). This is in compliance with the Statistics and Registration Service Act 2007.

v. The views of the public are considered in light of the data used and the perceived benefits of the research.

Whilst the views of the public have not been sought with regards to the research, there is a clear public benefit for the analysis (as described in section 1.). A public consultation on the Approved Researcher scheme in 2015/16 recommended that commercial organisations should be allowed to access ONS microdata where there is a clear public benefit.

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

To help promote greater transparency and in compliance with the updated Approved Researcher scheme, the researchers have agreed to their details being included on a public record of Approved Researchers and to publishing the findings of their research (including on the ONS Approved Researcher pages). The research results will be available to the public and the results will be presented to a number of policy organisations including the Greater London Authority and Home Office.

B7 How will the findings of the research be disseminated?

The results will be published by London First on their <u>website</u> within four weeks of the agreed final report. A link to the report will also be included on the ONS Approved researcher web pages in accordance with the Approved Researcher transparency criteria.

B8	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.
None	
	Section C

Details of Data Subjects

C1	Data subjects to be studied		
the	es the Study include all subsections of population (i.e. all ages, sex, ethnic ups etc	Yes No If no please detail which subsections with justification(s) below	
Sub	sections of the population (including vuln	erable groups) the project focuses on:	
The project focuses on migrant workers. The project will consider a range of categories such as industry sector, employment % and type (e.g. full time), salary band, and country of origin to help understand any variations in the impact on these from migrant workers.			
Dependents of migrants will be counted, but not analysed. For instance, PWC are interested that a migrant worker has two dependents, but they will not look at, or analyse, the characteristics of them.			
Anyone not economically active (of any age), or above working age, will be flagged as such and not included in the analysis.			
Justification for focusing on these subsections or groups:			
To meet the scope of the project and inform government policy in the contribution and impact of migrant workers on the London economy.			

NRS



C2 Please detail consent given to use data specified in section B2

No explicit consent gained so the ONS Approved Researcher gateway is being used.

C3 If you are using data held by a third party please detail how you will obtain this

N/A

Annex B: Summary of discussions on Microdata Release Panel referred application from PriceWaterhouseCoopers (PWC) request to use ONS data through the Approved Researcher Scheme

Considered via correspondence 22/11/2016-28/11/2016

Participating Members

Mr Ian Cope (Chair)
Mr Robert Bumpstead (Deputy Chair)
Ms Vanessa Cuthill
Mr Keith Dugmore
Mr Colin Godbold
Ms Annie Hitchman
Mr Neil McIvor
Ms Isabel Nisbet
Ms Marion Oswald
Mr Hetan Shah

UK Statistics Authority

Mr Adil Deedat

Office for National Statistics

Mr Pete Stokes (Microdata Release Panel)

Apologies:

Dr Dean Machin Mr Osama Rahman Professor Martin Severs

1. Commercial request for data

- 1.1 In early November PriceWaterhouseCoopers (PWC) made a request to access ONS data through the Approved Researcher scheme. This was the first request from a commercial organisation following changes to the Approved Researcher scheme which now allows commercial access to ONS data within secure environments, so long as the research proposal has a clear public benefit.
- 1.2 As agreed at the meeting of NSDEC on 5 July 2016 commercial requests to use ONS data within the Virtual Microdata Laboratory (VML) are considered by the Microdata Release Panel (MRP), which governs access to the VML, and are then referred to the National Statistician's Data Ethics Advisory Committee (NSDEC) for ethical consideration.
- 1.3 Ordinarily such requests would be considered at one of NSDEC's quarterly meetings. However in order for the proposal to deliver public benefits outlined, expedited review was in this instance required to ensure timely access to the data.
- 1.4 Members of NSDEC were provided with the project application and asked to consider the proposal, which had been approved by the MRP, via correspondence.

2. User Need and Public Benefit

2.1 There was agreement amongst members that the proposal had a clear user need and was of public benefit. Members also saw the need for an expedited review given the need for the research to be undertaken before triggering of article 50.

3. Data, Confidentiality and security

- 3.1 Members were in agreement that the safeguards around settings, training and disclosure were robust.
- 3.2 Queries were raised as to whether only data relevant to the proposal would be made available to researchers.
- 3.3 Members heard that researchers would have access to the complete de-identified datasets but would only use relevant information in their analysis. This is standard practice in all secure environments, as extensive security controls remove the need to create bespoke extracts.
- 3.4 Members sought assurance that no data or draft reports would be shared outside of the PWC team other than the final published report which will be shared with London First.
- 3.5 Mr Stokes informed members that only the named approved researchers would have access to the data and no report can be shared with anyone else including London first until the VML team have confirmed that those outputs are relevant to the research goals and are completely non-disclosive. Once analysis has been completed and outputs released following disclosure control, the outputs are subject to the disclosure provisions of the Statistics and Registration Service Act. These could then be discussed with others during compilation of the report. This normally happens in academia however researchers have stated that they will not do so in this instance.
- 3.6 It was suggested that the application should make this clearer.
- 3.7 Members sought clarity on the fate of the data used by the researchers once the report is completed.
- 3.8 Mr Stokes informed members that data never leave the VML, as VML security controls requires all analysis to take place within the secure environment. Any coding and analyses produced by the researchers will be archived securely within the VML in case it is needed in the future, for example to replicate, validate or challenge the eventual published results.
- 3.9 There was considerable discussion around the need for data on "number of dependents of migrants", "employment profile of migrant dependents" and inclusion of "individuals aged 70 and over".
- 3.10 It was suggested that more information on why access to information listed in 3.9 is relevant to the study. As the proposal is about the effects of migrant workers on the economy, inclusion of dependents should only relate to adult dependents as opposed to children. The application makes clear that only those aged 18 and over will be included in the study. Therefore, if the proposal requires inclusion of minors who are dependents of migrants then the application will need to be amended and the committee will need to reconsider the proposal.
- 3.11 Mr Stokes agreed to clarify these points with the researchers and report back to members. He also provided reassurance that no outputs will be released if they are not relevant to the agreed research goals and or present any risk to the confidentiality of any data subject.

3.12 It was also suggested that the application also makes clear that location information to be used in the analysis will be based on work place location rather than home residence.

Action: Mr Stokes to clarify 3.9 with the researchers and update NSDEC and amend the proposal as required.

4. Legal queries

- 4.1 Members sought clarity around ownership of the data to be used in the proposal, and that the research framework had been agreed by ONS and was deemed analytically sound.
- 4.2 Assurance was provided that that all datasets to be used in the study were owned by ONS and that ONS had had sight of the research framework and agreed that methods were of sufficient merit. There was however no requirement for PWC to share their research methods, however Approved researchers are encouraged to be open and collaborative. This is achieved through a public register of researchers and their projects which enables future researchers or readers of project reports to identify who they may want to contact to discuss their research in more detail. PWC has agreed to inclusion on this record and support its intention.
- 4.3 Additional clarity was sought on the legal gateway to access the data.
- 4.4 Members heard that the Microdata Release Panel, which considers all requests to use the VML includes members from ONS legal services, who confirmed that the Approved Researcher gateway is appropriate for the purposes of the study. In addition all proposed researchers on this project have already attained approved researcher status as individuals through demonstrating their qualifications and experience and by completing pan government training. They are therefore eligible to use data in the VML once project approval is given.

Report Publications

- 5.1 It was suggested that the entire report should be published to mitigate London first publishing only a redacted version. The full report should be published within four weeks of submission from London first.
- 5.2 Mr Stokes informed members that subject to project approval, a publication deadline will be agreed with researchers before analysis starts. This is standard practice. The MRP will also insist that publication is prompt. Member's proposal of four weeks after submission is consistent with timescales previously discussed with PWC and MRP are content in making this deadline a condition of access.
- 5.3 Members sought clarification on whether ONS would clear the project to ensure that ONS data is presented in a factual and unbiased way.
- 5.4 Members heard that ONS cannot make judgements on the validity of research undertaken using its data as doing so could compromise independence. If the research were published in an incorrect or misleading way then ONS would respond publically, as it does when ONS's open data are misinterpreted.

Conclusion

- 6.1 The Microdata Release Panel approved this project and referred to NSDEC for ethical consideration via correspondence.
- 6.2 All participating members agreed that the project demonstrated clear public benefit.

6.3 Members agreed that the project could proceed subject to use of data as detailed in 3.9 being made clearer and reflected in the application. Other conditions include the full report being published within four weeks of the report being made available.

Adil Deedat, NSDEC Secretariat, Central Policy Secretariat, UK Statistics Authority, 28 November 2016

UK Statistics Authority

National Statistician's Data Ethics Advisory Committee

NSDEC(17)02

Applications: policies and process for expedited review

Purpose

1. This paper outlines options for NSDEC to review proposals which require expedited ethical review outside of its quarterly meetings.

Recommendations

- 2. Members of NSDEC are invited to:
 - consider the options outlined for reviewing proposals (at 21) between meetings; and
 - agree an option as to how NSDEC should consider proposals which require ethical review in between meetings.

Background

- At the meeting of NSDEC on 14 October 2015 members agreed processes for ethical review [NSDEC(15)06]. Members also discussed the possibility of forming a subcommittee should demand and/or logistics require ethical review of proposals between meetings.
- 4. Whilst NSDEC's quarterly meetings have generally been, and continue to be sufficient to meet demand of those requiring ethical review, there are now likely to be an increase in cases where an expedited ethical review is required. This is in part due to NSDEC's role in providing ethical review of ADRN projects and also as a result of NSDEC's more recent role in providing ethical review of projects referred to NSDEC by the Microdata Release Panel.
- 5. This paper therefore outlines options for accommodating such proposals between meetings of NSDEC.

Discussion

 Here we illustrate examples of when an expedited review may be required and provide a number of options for NSDEC to accommodate proposals outside of its quarterly meetings.

Referral from the MRP

- 7. NSDEC recently agreed to review proposals from commercial and other organisations applying to use ONS data through the Approved Researcher scheme. These requests are referred to NSDEC for ethical consideration by the Mircrodata Release Panel (MRP) which governs access to the Virtual Microdata Laboratory (VML).
- 8. In November 2016, the MRP received its first request from a commercial body to access ONS data held in the VML. The MRP approved this application to use ONS data but referred it to NSDEC for ethical consideration.
- 9. Ordinarily, this application would have been considered at one of NSDEC's quarterly meetings. This proposal, however, required analysis to be undertaken and findings reported ahead of NSDEC's next meeting in order for user needs to be met and public benefits to be realised.
- 10. As commercial organisations are now eligible to apply to use ONS data through the Approved Researcher scheme, and given that these proposals must undergo ethical review before access is granted, NSDEC is likely to receive similar requests moving forward.

- 11. Existing provisions within the Approved Researcher scheme allow academics to access ONS data following approval of their project by the Microdata Release Panel (ethical review is not mandatory for academic projects although often a review has been undertaken by their host institution). The additional requirement for commercial organisations to gain ethical review from NSDEC means that commercial access to data may be less timely than requests originating from academia.
- 12. In order to ensure commercial researchers gain as timely access to data as those from academia (who do not require ethical review), consideration needs to be given as to how NSDEC can accommodate such proposals outside of its quarterly meetings, to ensure access to data is not hindered.
- 13. Where proposals are referred by the MRP to NSDEC, Mr Peter Stokes, member of the MRP, will represent the application on behalf of researchers.
- 14. Whilst in the first instance, demand from the MRP is likely to be high, over time consideration of such proposals is likely to be more stream lined as NSDEC sets precedents.

Administrative Data Research Network

- 15. At the last meeting of NSDEC on 11 October, members reviewed its first two ADRN projects from Welsh Government. Members advised major revisions to one of these proposals and agreed that once addressed to re-consider the proposal via correspondence.
- 16. ADRN projects require approval from the ADRN Approvals Panel which meets monthly. In order to ensure ADRN projects can proceed as quickly as possible there may be circumstances where ADRN projects are considered between meetings.
- 17. In light of this NSDEC may wish to define under what circumstances and how projects from the ADRN or otherwise are considered or re-considered in between meetings.

Scope for expedited review

- 18. Above we have illustrated examples of the types of proposals which could require an expedited ethical review.
- 19. It is envisaged that in the first instance expedited review should be limited to either ADRN proposals or those originating from the Microdata Release Panel.
- 20. In both these instances there must also be a clear reason for requiring an expedited ethical review. For example this may be:
 - i. where the benefits of the proposal to use ONS data through the approved researcher scheme cannot be realised by waiting for the next NSDEC meeting, as would have been the case for [NSDEC17(01)];
 - ii. where waiting for the next meeting means timeliness of access to ONS data through the approved research process, is hindered.
 - iii. where ADRN projects have been approved by the approvals panel (which meets monthly) shortly after an NSDEC meeting; or
 - iv. where an ADRN project has been approved by the approvals panel but received major revisions at a previous NSDEC meeting.
- 21. In all other instances, for example proposals originating from the Office for National Statistics, Devolved Administrations or the Government Statistical Service, the secretariat will continue to work with officials to ensure these are planned accordingly and are considered at quarterly meetings.

Options for accommodating proposals requiring expedited review

- 22. There are a number of options by which NSDEC could review proposals requiring expedited review or reconsideration:
 - i. **Option 1:** all members to consider proposals by correspondence, telephone or additional face to face meetings; or
 - ii. **Option 2:** form a sub-committee to consider proposals by correspondence, telephone or face to face meetings. Any subcommittee would need to represent a quorate meeting, and therefore consist of four members with at least two independent members; or
 - iii. **Option 3:** increase the frequency of NSDEC meetings but reduce the length of meetings.
- 23. To ensure that expedited reviews are valid, in choosing an option, NSDEC may wish to consider the number of members who need to participate in an expedited review.
- 24. Under the existing terms of reference, meetings (and therefore NSDEC's decisions) are quorate where four or more members are present including at least two independent members as well as the Chair or the Chair's delegated nominee.
- 25. To ensure NSDEC maintains its commitment to transparency, where there has been ethical consideration of a proposal outside of NSDEC's quarterly meetings, the application and summary of comments will be reviewed and published as part of NSDEC's next meeting, as is presented in this pack for [NSDEC17(01)].

Adil Deedat, NSDEC Secretariat, Central Policy Secretariat, UK Statistics Authority, 5 January 2017

UK Statistics Authority

National Statistician's Data Ethics Advisory Committee

NSDEC(17)03

MRP: Data analysis to inform research on the size, scope, dynamics and geographic clusters of digital tech industry activity in the UK

Purpose

 This paper presents a proposal referred by the Microdata Release Panel for access and use of ONS data by Frontier Economics who have been commissioned by Tech UK to analyse ONS business data to understand tech industry activity in the UK.

Recommendations

- 2. Members of NSDEC are invited to consider the 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 it be stopped from proceeding.

Background

- 3. In 2010, the former Prime Minister David Cameron launched Tech UK to support East London's tech cluster. The organisation has since grown to support Greater London, Manchester and other major cities in the UK.
- 4. In 2016 Tech UK published a report on the size, scope, dynamics and geographic clusters of digital tech industry activity.
- 5. Tech City has now commissioned Frontier Economics to facilitate the collection and tabulation of specific business sector data as part of preparing report "Tech Nation 2017".
- 6. The proposal would see Frontier Economics use ONS's Business Structure Database for 1997 to 2015 inclusive and the Annual Business Survey for all years from 2008.
- 7. Analysis will be undertaken within the ONS secure Virtual Microdata Laboratory.
- 8. Digital tech businesses are transforming the employment landscape, driving productivity and reimagining traditional industries.
- 9. This comprehensive project will research the growth of digital tech clusters to gauge the UK's Digital Tech Economy and the wider impact on business, employment and economic trends.
- 10. The detailed findings will be beneficial to investors, collaborators, customers, educators and policymakers and will help promote greater understanding of these tech businesses and inform policy and decision making.

Adil Deedat, NSDEC Secretariat, Central Policy Secretariat, UK Statistics Authority, 3 January 2017

List of Annexes

Annex A: Application: MRP: Data analysis to inform research on the size, scope, dynamics and geographic clusters of digital tech industry activity in the UK, Mr Peter Stokes, Office for National Statistics, 12 January 2017

9



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 <u>plain English</u> 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



Section A Application Details

A1 Responsible Owner		
Full	Name:	Position: Head of Data Access and Exploitation
Add	ress	Email:
		Organisation: ONS

Declaration to be signed by the responsible owner

I have met with and advised the applicant on the ethical aspects of this project design (applicable only if the responsible owner is not the Applicant).

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.

I am satisfied that the research complies with current professional, departmental and other relevant guidelines.

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.

I will provide notification when the study is complete if it or fails to start or is abandoned.

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.

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	rint	Name:
_		

Signature:

Date January 2017



A2 Applicant Details (if applicant is not the responsible owner)		
Full Name	Position	
Address:	Email:	
	Telephone:	
	Organisation: Frontier Economics Ltd	
	1	

Project Title: Data analysis to inform geographic clusters of digital tech income.	rm research on the size, scope, dynamics and dustry activity in the UK.
Start Date: 1 February 2017	End Date: 31 March 2017
Project Sponsor (select all that apply)	
□ ons	
ADRN	
□ GSS	
☐ Collaboration	
Other (Please specify)Tech City U	K http://www.techcityuk.com/

ist of Collaborators/Sponsors	Details and relevant documentation relating to collaboration (you may attach
	copies of relevant documentation)

Policies

A5 Proposed Site of Research (select all that apply)
Where will the research take place?
ONS
▼ ∨ML
☐ 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)

Tech City UK was launched by the UK Prime Minister in 2010 to support the East London tech cluster. The organisation has grown to support Greater London, Manchester and other major cities around the UK.

They do this in three ways:

- Targeted programmes: that fill market gaps across the whole lifecycle of Digital Business Academy to Future Fifty
- Engaging policing convening: providing an agile and responsive feedback loop to government to help shape the right policies via roundtables, surveys, white papers



and reports such as Tech Nation

 Effective promotion and thought leadership: championing and connecting the digital sector in London, the UK, and internationally via trade missions, media engagement and events.

Tech City UK plan to update earlier research (published in 2016) on the size, scope, dynamics and geographic clusters of digital tech industry activity in the UK. The annual <u>Tech Nation report</u> describes the trends in the digital economy and is published on the Tech City website. The 2016 report included a foreword from the then UK Prime Minister, David Cameron. A similar MRP research project carried out by Frontier Economics (on behalf of Tech City UK) was approved in 2015 for the preparation of the 2016 report.

Tech City has commissioned Frontier Economics to assist with the collection, tabulation and analysis of specific business sector datasets for the preparation of the report, "Tech Nation 2017". This will report on the geography of the digital tech industries in the UK, contributing to the broad mandate of Tech City UK from the Department of Culture, Media and Sport. The report will be published in February 2017, and will include interactive data visualisations based on the results. The findings from the research will also inform Tech City's efforts to support and showcase the digital tech industries in the UK.

The project will make use of industry big data and official data sources, including ONS Business Structure Database (BSD) and Annual Business Survey (ABS), to estimate economic statistics about the Digital Tech Economy. The data sources will be analysed to identify the capabilities, resources and infrastructures of tech clusters across the UK. The project will include an online survey of digital tech businesses and qualitative interviews with stakeholders.

Frontier will draw on a combination of matched administrative (BSD) and survey datasets (ABS) in the ONS Virtual Microdata Laboratory (VML) secure environment to construct size estimates and cluster analysis (location quotients, Kernel estimation) of industry and sub-industry statistical aggregations of business counts, employment, turnover, and Gross Value Added (GVA), at various levels of geography for the UK. The aggregated data tabulations will be used together with additional information collected from a separate online survey of, and qualitative interviews with, digital tech companies. The final results will be presented in a very comprehensive Tech Nation 2017 report which provides a high profile assessment of the digital technology sector in the UK each year. The report serves as a statistical reference guide on the digital technology sectors and is used by a range of stakeholders, including central and local government policy makers, to develop digital tech strategies.

The research project was approved by the Microdata Release Panel (MRP) on 16 December 2016 on the grounds that there was a legal gateway to access the data, it was appropriate use of ONS data and a public benefit was demonstrated (such as an improved understanding of the characteristics of the creative and digital technology industry activity in the UK).

As this MRP project being carried out by a commercial organisation on behalf of another commercial organisation, the project is (as agreed) being referred to NSDEC to assess the ethical use of these data by a commercial organisation. If the project were being carried out



by a non-commercial organisation, it would have been agreed without NSDEC consideration. Frontier Economics have carried out a number of research projects in the VML, including on behalf of the Department of Business, Energy and Industrial Strategy (BEIS) and the Department of the Environment, Food and Rural Affairs (DEFRA).

The urgency of the research is determined by the need to complete the analysis in time for publication of Tech Nation 2017 in February this year.

B2	Data Use
152 1	Data 036

Type of data	Data Level Please specify the name of the data set			
	Aggregat e Data	Identifiable Data	De- identified personal data	Anonymised/ pseudo anonymised
Administrative data (please specify, e.g. Patient Register 2011, School Census 2012 etc, in the relevant options adjacent)				
Big Data (please specify e.g. Twitter data, smart meters and mobile phones, in the relevant options adjacent)				
Survey Data (please specify e.g.LFS, BRES, etc in the relevant options adjacent)			Business Structure Database (BSD), all years 1997 – 2015. Annual Business Survey (ABS), all years from 2008	
Census Data (please specify year, e.g. Census 2011 in the relevant options adjacent)				
Other (please specify e.g. Ordinance Survey Address register in the relevant options adjacent)				

Policies



B3 How will information be kept confidential and data kept secure? (max 500 words)

Access to data will only take place within the secure VML environment and all outputs will be checked by the VML team prior to release to ensure disclosure control and the confidentiality of data subjects is protected. All analysis and use of the data will be within the VML at a secure VML setting at one of the ONS offices. The Frontier Economics researchers have been accredited as Approved Researchers and have successfully completed a number of projects using data in the VML. These were approved prior to the establishment of the National Statistician Data Ethics Advisory Committee. They have successfully completed the training on the safe use of data in research environments (SURE). No data or report drafts will be shared with anyone outside of the project team.

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

Frontier Economics will use a combination of administrative (Business Structure Database, BSD) and survey datasets (Annual Business Survey, ABS) in the VML to develop estimates on the size and geographic location of industries and sub-industries for the UK's "digital tech" industries (as classified by Standard Industry Classification (SIC) codes). They will produce location quotients (LQ) to quantify how concentrated the high-tech industries are in a region as compared to the nation. This should reveal what makes a particular region "unique" in comparison to the national average.

Specifically, Frontier will produce aggregations of business counts, employment turnover and Gross Value Added (GVA) estimates (to measure the value of the goods and services produced) for these high tech industries at various levels of geography (countries, regions and Travel to Work Areas - TTWAs). The BSD includes SIC, location, employment and turnover data for all UK businesses registered for PAYE/VAT. The ABS has 2007 SIC, location and detailed financial data which allows the estimation of approximate GVA figures. Much like how ONS regularly produce publicly available aggregates for these same measures at less granular levels of industry and geography combinations, the main objective of the report is to produce "custom" statistical aggregations for the size and geography of this group of industries.

The project will also add two new bits of analysis—generating tabulations of firm entry and exit, and looking at so-called high-growth firms which may be defined as either falling within a certain percentile growth distribution or achieving certain growth targets over a defined



period of time. It follows a very similar methodology to the MRP research project (ref 1000439) carried out in the VML in 2015, which culminated in the preparation of the Tech Nation 2016 report. Last year's project also included analysis for the creative industries (this year's project will not). Frontier Economics very much welcome the Committee's suggestion to explore the use of Company's House data and employing methods such as text mining, and they will follow these up for this project and similar research in the future.

In summary, the main outputs from the VML will be aggregate counts of business units, employment, and turnover by year (2007-2015), geography (country, region, TTWA), and sector ("digital tech" versus all others), as well as average GVA per worker for limited years and geographies. It will also include similar counts for businesses fitting certain criteria — those newly born, those closing doors, those achieving various measures of "high growth" (e.g. the number of firms that either (a) are in the top 10 percent of the one-year growth distribution for employment or turnover, or (b) achieved at least 78% growth in employment or turnover over a three-year period).

The data will be used to show trends in levels, growth, and geographic concentration of the digital tech sectors throughout the UK. The findings will be published in the high-profile Tech Nation report, and made available for public consumption and use. The data produced in the VML will be part of a larger research product in the Tech Nation report, including data on job openings and other information that has been gathered from web resources.

B5 Please outline the proposed benefits of the project (max 500 words)

Digital tech businesses are at the heart of the UK economy and are playing an important role in driving growth. The impact of this dynamic sector is profound, predicated on a fundamental belief in innovation and doing things differently. Digital tech businesses are transforming the employment landscape, driving productivity and reimagining traditional industries. It is estimated that they contribute some £87 billion to the economy, and providing security and job opportunities to local communities. The 2016 Tech Nation report highlighted that the growth isn't focused on hubs such as London or Leeds, but spread across the country.

This comprehensive project will research the growth of digital tech clusters to gauge the UK's Digital Tech Economy and the wider impact on business, employment and economic trends. The detailed findings will be beneficial to investors, collaborators, customers, educators and policymakers and will help promote greater understanding of these tech businesses.

The Government is behind the country's digital transformation, backing new technologies, investing in infrastructure, supporting investment, removing barriers to innovation and helping upgrade the digital skills that a modern work force needs. The report, which will be published in February 2017, and the research findings will provide insight of the contribution that the digital technology is making to employment in digital and traditional industries, and

Policies

9

NRS

 ∞



to the economy across the country. The report will provide a time series analysis to help benchmark changes to, and provide greater understanding of, the UK's digital economy. The findings will also inform Tech City's efforts to support and showcase the digital tech industries in the UK. This will deliver a public benefit in terms of providing an evidence base for public policy decision and service planning to help stimulate economic growth and create the right infrastructure to help these businesses thrive. The work is also likely to benefit our understanding of the digital economy and therefore help contribute to delivery of the Bean Review recommendations.

The 2016 Tech Nation reported:

- The UK's Digital Tech Industries are growing 32% faster than the rest of the UK economy (in turnover)
- The estimated turnover of UK digital tech industries reached £161bn in 2014
- The Digital Tech Industries are creating employment opportunities and accounting for 1.56M jobs across the UK
- The UK's increasingly diverse range of digital skills and specialisms are also now thriving beyond the tech sector – 41% of Digital Tech Economy jobs exist in what are thought of as traditionally non-digital industries – such as the public sector and financial services
- Over 80% of Tech Nation clusters have seen growth in digital turnover, jobs and advertised digital salaries
- -See more at: hhtp://www.techcityuk/com/technation/

The report received considerable media coverage, including <u>The Guardian</u> and the <u>Telegraph</u>.

There will be additional public benefits from the research, including to benchmark with the findings from the 2016 and 2015 Tech Nation reports which should help to significantly extend understanding of social or economic trends or events by improving knowledge or challenging widely accepted analyses.

Policies



Please outline the ethical issues that might arise from the proposed study and how they will be addressed (all research projects have some ethical considerations, so this section must not be left blank)

i. The use of data has clear benefits for users and serves the public good

The research findings will provide insight of the contribution that the digital technology is making to employment in digital and traditional industries, and to the economy across the country. It will inform the 2017 Tech Nation report, which will be published in February 2017, and provides a time series to help benchmark changes to, and provide greater understanding of, the UK's digital economy. The findings will also inform Tech City's efforts to support and showcase the digital tech industries in the UK. This will deliver a public benefit in terms of providing an evidence base for public policy decision making and to provide an evidence base for decisions which are likely to significantly benefit the UK economy, society or quality of life of people in the UK. The Government supports the production of the Tech Nation report and the Prime Minister at the time, David Cameron provided the foreword for last year's report.

The 2016 Tech Nation report received media coverage from <u>The Guardian</u> and the <u>Telegraph</u> which supported public debate. The Guardian reported that salaries in the digital tech businesses were on average £15,000 higher than non-digital companies and in response Alexandra Jones, the chief executive of the think tank, Centre for Cities, called for more work to be done to tackle congestion and a shortage of affordable homes in areas that house a large number of digital companies.

ii. The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.

The data will be accessed in the secure VML accredited environment and is deidentified to protect the confidentiality of data subjects. The Approved Researcher
scheme will be the legal gateway used to access the data and the researchers have
successfully completed the Safe User of Research data Environments (SURE) training
(and passed the assessment) and have signed to confirm they adhere to the Approved
Researcher declaration setting out how they will manage the data and protect the
confidentiality of data subjects in line with the Statistics and Registration Service Act
2007. The researchers will not be able to take the data outside of the VML and their
outputs – both intermediate and final - will be checked and cleared by the ONS VML
branch to ensure the confidentiality of data subjects is protected.

The report will combine statistical summaries and data visualisations in aggregate form, and correlation analysis of data to reach firm conclusions on the data. The final report will be fact-based so that any conclusions are based on the data, rather than assumptions.

iii. 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.

No new technologies are being used. The research methods employed will be openly available for further scrutiny or replication of results.

iv. 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.

Access to the potentially disclosive data will be in a secure environment (VML) and via an approved legal gateway (Approved Researcher scheme). This is in compliance with the Statistics and Registration Service Act 2007.

v. The views of the public are considered in light of the data used and the perceived benefits of the research.

Whilst the views of the public have not been sought with regards to the research, there is a clear public benefit for the analysis (as described in section 1.). A public consultation on the Approved Researcher scheme in 2015/16 recommended that commercial organisations should be allowed to access ONS microdata where there is a clear public benefit.

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

Tech City UK will publish a comprehensive report (Tech Nation) with aggregated tables that are not disclosive, on its website. A link to the report will be included on the ONS website. The previous year's research findings will also be available and compared, to help benchmarking. The report will be an entirely factual analysis (so not political, or opining in the value of creative tech industries). All analysis will be carried out in the VML and all outputs will be checked to ensure the confidentiality of data subjects is protected.

To help promote greater transparency and in compliance with the updated Approved Researcher scheme, the researchers have agreed to their details being included on a public record of Approved Researchers and to publishing the findings of their research (including on the ONS Approved Researcher pages).

B7 How will the findings of the research be disseminated?

The results will be published by Tech City UK on their <u>website</u>. Last year's Tech Nation report is published on the website as well as other reports that have made use of published ONS data, such as the London 2036: An agenda for growth and jobs. A link to the Tech Nation report will also be included on the ONS Approved researcher web pages in accordance with the Approved Researcher transparency criteria.

	B8	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.	
•	None		

Section C Details of Data Subjects

C1	Data subjects to be studied		
Does the Study include all subsections of the population (i.e. all ages, sex, ethnic groups etc ☐ Yes No ☐ If no please detail which subsections with justification(s) below			
Sub	sections of the population (including vulne	erable groups) the project focuses on:	
The project focuses on digital tech industry activity in the UK. It will consider the industry and sub-industry statistical aggregations of business counts, employment, turnover and GVA, at various levels of geography for the UK.			
Justification for focusing on these subsections or groups:			
To meet the scope of the project and gain an understanding of the size, scope, dynamics and geographical clusters of the digital tech industry activity in the UK. This will inform government policy and planning on the contribution and impact of these industries on the economy.			

9

NRS

C2 Please detail consent given to use data specified in section B2

No explicit consent gained so the ONS Approved Researcher gateway is being used.

C3 If you are using data held by a third party please detail how you will obtain this

Not applicable.

Policies

UK Statistics Authority

National Statistician's Data Ethics Advisory Committee

NSDEC(17)04

Microdata Release Panel: Quantifying the extent and variety of incidents involving suicide on England's strategic road network

Purpose

This paper presents a proposal for use of ONS data by Road Safety Analysis Ltd (RSA)
who have been commissioned by Highways England, to quantify the extent and variety
of incidents involving suicide on England's strategic road network.

Recommendations

- 2. Members of NSDEC are invited to consider the 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 it be stopped from proceeding.

Background

- 3. Highways England, a government agency within the Department for Transport are responsible for operating, maintaining and improving England's motorways and major A roads.
- 4. Highways England has identified eight research priorities and delivery of these will contribute to its four year strategy. RSA has been commissioned to undertake analysis on its behalf.
- 5. The purpose of the project is to:
 - i. identify and define future actions which may prevent such incidents from occurring:
 - ii. protect road users from collateral harm; and
 - iii. manage consequent disruption.
- 6. ONS will make its mortality data from 2001 to 2015 available to RSA researchers within the secure ONS Virtual Microdata Laboratory.
- 7. The research outcomes will achieve a public benefit by providing an evidence base for public policy decision making and public service delivery by helping to prevent injury and deaths on the roads in terms of public health (e.g. support for vulnerable users and workers), and design and engineering solutions on the road network.
- 8. The research outcomes will contribute directly to Highways England's four year research strategy and its priorities, including improvements to the health safety of road users and workers.

Adil Deedat, NSDEC Secretariat, Central Policy Secretariat, UK Statistics Authority, 3
January 2017

List of Annexes

Annex A: Application: Quantifying the extent and variety of incidents involving suicide on England's strategic road network, Mr Pete Stokes, Microdata Release Panel, Office for National Statistics, 16 January 2017



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

Section A Application Details

A1 Responsible Owner	
Full Name:	Position: Head of Data Access and Exploitation
Address	Email:
	Telephone:
	Organisation: ONS

Declaration to be signed by the responsible owner

I have met with and advised the applicant on the ethical aspects of this project design (applicable only if the responsible owner is not the Applicant).

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.

I am satisfied that the research complies with current professional, departmental and other relevant guidelines.

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.

I will provide notification when the study is complete if it or fails to start or is abandoned.

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.

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.

Print	Name:
Signa	ature:

Date January 2017



Full Name	Position
Address:	Email:
	Telephone:
	Organisation: Road Safety Analysis Ltd (RSA)

A3 Project Information	
Project Title: Quantify the extent and varied strategic road network.	ty of incidents involving suicide on England's
Start Date: 1 February 2017	End Date 21 October 2017
Project Sponsor (select all that apply)	
□ ons	
□ ADRN	
□ GSS	
☐ Collaboration	
Other (Please specify)Highways England https://ww	ww.gov.uk/government/organisations/highways-england

List of Collaborators/Sponsors	Details and relevant documentation relating to collaboration (you may attach copies of relevant documentation)
Highways England	



A5 Proposed Site of Research (select all that apply)
Ad 1 Toposed Oite of Nesearch (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

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

Project Details

Highways England is the government organisation within the Department for Transport charged with operating, maintaining and improving England's motorways and major A roads. Formerly the Highways Agency, it became a government organisation/body in April 2015.

Highways England does not manage all roads in Britain.

- Local roads are managed by the relevant local authority;
- Scottish roads are managed by Transport Scotland;
- Welsh roads are managed by the Welsh Assembly; and



London roads are managed by Transport for London.

Traffic and transport legislation, regulations and policy are the remit of the Department of Transport and the research outcomes will inform their policy work.

Highways England has identified eight research priorities and delivery of these will contribute to its four year research strategy. The priorities include improvements to road user and workforce health and safety.

Road Safety Analysis (RSA) Ltd has been commissioned by Highways England to quantify the extent and variety of incidents involving suicide on England's strategic road network. This is part of as wider portfolio of reports examining personal risk on England's strategic roads and will contribute to Highway England's research strategy.

The purpose of the project is to:

- identify and define future actions which may prevent such incidents from occurring;
- · protect road users from collateral harm; and
- manage consequent disruption.

To deliver these aims, RSA will:

- examine in detail the locations and timings of road related suicide incidents;
- identify the negative consequences for other road users arising from the incidents;
- attempt to match the incidents with other datasets where possible;
- analyse relevant demographic trends to help understand whether particular societal groups are at disproportionate risk; and
- investigate the viability of future multi-agency co-operation to address these and related issues.

Because confirmed suicides on public roads do not count as road casualties that are reported to the police, and subsequently reported using the STATS19 accident reporting form, ONS data is necessary in order to assess this aspect of road safety more effectively. Perhaps as a consequence of the challenges of collecting the data, a detailed examination of road related incidents involving suicides has not been carried out. There may be significant outcomes for both the road safety planning and public health communities more generally, in gaining a more complete understanding of how the road network impacts on suicide events and vice versa, as well as enhancing opportunities for multi-agency collaboration to manage and reduce risks.

The output from the research will be a high level summary report elucidating trends in suicide incidents on the strategic road network, which will be supplied to Highways England. The report will include data tables, thematic maps and charts to illustrate trends, but will not include any detail of individual incidents or any other information from which persons could be identified. A link to the report will be included on the ONS website.

The Microdata Release Panel (MRP) approved the proposal at its meeting on 16 December 2016 on the grounds that there was a legal gateway to access the data, it was appropriate use of ONS data and a public benefit was demonstrated (provides an evidence basis for



public health and transport decision making and service planning).

B2	Data	Use
----	------	-----

Type of data	Data Level Please specify the name of the data set			
	Aggregat e Data	Identifiable Data	De- identified personal data	Anonymised/ pseudo anonymised
Administrative data (please specify, e.g. Patient Register 2011, School Census 2012 etc, in the relevant options adjacent)				
Big Data (please specify e.g. Twitter data, smart meters and mobile phones, in the relevant options adjacent)				
Survey Data (please specify e.g.LFS, BRES, etc in the relevant options adjacent)			ONS Mortality data, 2001 – 2015.	
Census Data (please specify year, e.g. Census 2011 in the relevant options adjacent)				
Other (please specify e.g. Ordinance Survey Address register in the relevant options adjacent)				

B3 How will information be kept confidential and data kept secure? (max 500 words)

Access to data will only take place within the secure VML environment and all outputs will be checked by the VML team prior to release to ensure disclosure control and the confidentiality



of data subjects is protected. All analysis and use of the data will be within the VML at a secure VML setting at one of the ONS offices. The RSA have applied to be accredited as Approved Researchers and will complete their training on safe use of data in research environments (SURE) course on 18 January 2017. No data or report drafts will be shared with anyone outside of the project team.

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

RSA will use ONS mortality data from 2001 – 2015. The requested data is for all deaths occurring between 2001 and 2015 with ICD-10 codes X80.4, X81.4, X82.4, Y30.4, Y31.4 or Y32.4. The 'X' codes listed indicate intentional self-harm causes of morbidity and mortality which relate to jumping or falling from a high place, or crashing of a motor vehicle. The 'Y' code identifies events of undetermined intent cause of morbidity and mortality. The '.4' code records the place where the incident occurred, i.e. road-related. Access to this code only ensures that the researcher only has access to suicides that occurred on the road network.

The variables requested are:

- Postcode of place of death
- Date of death
- Time of death
- Age
- Sex
- Whether person committed suicide was in charge of the vehicle, and if so, what type
- Any road infrastructure involved
- indication of other parties involved
- Home postcode at time of death
- Location information provided by coroners

The Home postcode at time of death has been requested for the following purposes:

- Calculating the distance between deceased individuals' homes and their place of death, to explore the extent to which different types of road suicide relate to local communities in the vicinity where they occur; and
- Identifying the types of community in which deceased individuals lived, both in terms
 of ONS measures such as deprivation and rurality and also socio-demographic
 classification, in order to provide insight into what kinds of people are most often at
 risk.



Postcode analysis would remain entirely within the VML, and would not form part of the research output. Rather, output would aggregate deceased individuals by relevant metrics: for example, "Suicides from South East England on England's trunk roads 2001 to 2015 categorised by IMD quintile", or "Average distance from home of car driver suicides on trunk roads in North East England 2001 to 2015".

Stage 1: Not all suicide incidents are road-related, so the first stage of the research will examine incident locations to identify whether they occurred on a road and if so, which authority was responsible for managing it.

Incidents which did not occur on a road or related infra-structure will be excluded from further research. Those that are road-related, will be categorised by highway authority that manages that specific road, e.g. Highways England or the relevant local authority. Incidents that occurred on roads managed by Highways England and its predecessors (e.g. Highways Agency), will be examined in detail.

Stage 2: The next research stage will attempt to relate suicide incidents which may be recorded in other datasets such as police reported incidents recorded on a STATS19 form or Highways England network incident records. Where this proves possible, additional information concerning the event, such as other casualties which may have resulted and any further impact on the network, will be recorded.

Information on location and involvement of other parties in the incident will be obtained from free text fields from coroner reports. This information will be provided and checked by ONS to ensure that it does not include any identifiable data on the deceased or other parties.

Stage 3: Finally, all available information will be synthesised at a high level to identify any significant trends in locations, dates, times, demographics, engineering features and/or involvement of other road users. Only the trends discerned from this process will be included in the final report to Highways England. No personal information about the individual deceased persons will be supplied.

The final outputs and any intermediate outputs will be cleared by ONS to ensure the confidentiality of data subjects is protected. The output from the research will be a high level summary report elucidating trends in suicide incidents on the strategic road network, which will be supplied internally to Highways England. The report will include data tables, thematic maps and charts to illustrate trends, but will not include any detail of individual incidents or any other information from which persons could be identified. A link to the report will be included on the ONS website.

B5 Please outline the proposed benefits of the project (max 500 words)



Road Safety Analysis will quantify the extent and variety of incidents involving suicide on England's strategic road network. This is part of as wider portfolio of Highways England's reports examining personal risk on England's strategic roads.

The research outcomes will:

- help develop a better understanding of the types of road incidents involving suicide;
- identify and define future actions which may prevent such incidents from occurring;
- inform thinking about solutions to protect road users from collateral harm; and
- inform decisions about managing consequent disruption.

These research outcomes should achieve a public benefit by providing an evidence base for public policy decision making and public service delivery by helping to prevent injury and deaths on the roads in terms of public health (e.g. support for vulnerable users and workers), and design and engineering solutions on the road network. The research outcomes will contribute directly to Highways England's four year research strategy and its priorities, including improvements to the health safety of road users and workers.



B6 Please outline the ethical issues that might arise from the proposed study and how they will be addressed (all research projects have some ethical considerations, so this section must not be left blank)

i. The use of data has clear benefits for users and serves the public good

Road Safety Analysis will quantify the extent and variety of incidents involving suicide on England's strategic road network. This is part of as wider portfolio of Highways England reports examining personal risk on England's strategic roads.

The research outcomes will:

- help develop a better understanding of the types of road incidents involving suicide;
- identify and define future actions which may prevent such incidents from occurring;
- inform thinking about solutions to protect road users from collateral harm; and
- inform decisions about managing consequent disruption.

The research outcomes will therefore achieve a public benefit by providing an evidence base for public policy decision making and public service delivery in relation to transport safety and public health. This will help to prevent further deaths on British roads and inform targeted interventions from public health organisations.

The Home postcode at time of death has been requested for the following purposes:

- Calculating the distance between deceased individuals' homes and their place of death, to explore the extent to which different types of road suicide relate to local communities in the vicinity where they occur; and
- Identifying the types of community in which deceased individuals lived, both in terms
 of ONS measures such as deprivation and rurality and also socio-demographic
 classification, in order to provide insight into what kinds of people are most often at
 risk.
- i. The data subject's identity (whether person or organisation) is protected, information is kept confidential and secure, and the issue of consent is considered appropriately.

The data will be accessed in the secure ONS Virtual Microdata Laboratory (VML) and is de-identified to protect the confidentiality of data subjects. The Approved Researcher scheme will be the legal gateway used to access the data and the researchers will have successfully completed the Safe User of Research data Environments (SURE) training and will sign and adhere to the Approved Researcher declaration setting out how they will manage the data and protect the confidentiality of data subjects in line with the Statistics and Registration Service Act 2007. The researchers will not be able to take the data outside of the VML and their outputs – both intermediate and final – will be checked and cleared by the ONS VML branch to ensure the confidentiality of data subjects is protected.

Policies



Because postcodes and ages are present in the raw data, there is potential for the data to be disclosive. Therefore threshold rules will be applied to all table cells in the output. Tables will summarise home postcode data geography by region to prevent disclosure. Data by postcode geography may be used for socio demographic analysis using metrics such as the index of multiple deprivation, but results will be summarised into national deciles.

A different geographical focus will be used for place of death, where information for complete routes will be calculated by road number (e.g. the M25). In this case, threshold rules will be used to prevent disclosure. Individual ages will not be used for summarising data – only ten year age bands will be considered. Chronological data will be presented as annual totals with no sub-divisions (e.g. quarters will not be used) or by month or hour of death as single totals calculated across all 15 years of the supplied data.

FSA Ltd have a signed memorandum with Highways England to further ensure they protect the confidentiality of data subjects.

iii. 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.

No new technologies are being used. The research methods employed will be openly available for further scrutiny or replication of results. RSA acknowledge the National Statistician's Data Ethics Advisory Committee suggestion that there is a risk of harm, albeit small, to the researchers given the sensitivity of the data. RSA will put arrangements in place to deal with any such possibilities.

RSA is aware that historic incident data for the network is available and can be used to obtain distribution and trends over time in a generalised way (e.g. to identify sites where attempted suicides have occurred). This data does not contain personal information about individuals and therefore will not be used for profiling at-risk individuals.

iv. 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.

Access to the potentially disclosive data will be in a secure environment (VML) and via an approved legal gateway (Approved Researcher scheme). This is in compliance with the Statistics and Registration Service Act 2007. As the Data protection Act does not apply to persons who are deceased, it is not applicable for the project.

v. The views of the public are considered in light of the data used and the perceived benefits of the research.

Whilst the views of the public have not been sought with regards to the research, there is a clear public benefit for the analysis (as described in section 1.). A public consultation on the Approved Researcher scheme in 2015/16 recommended that commercial organisations should be allowed to access ONS research data where there



is a clear public benefit.

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

The research methodology and outcomes will be made public and reported on as they contribute to Highway England's strategic research programme. A link to the research findings will be included on the ONS website.

To help promote greater transparency and in compliance with the updated Approved Researcher scheme, the researchers have agreed to their details being included on a public record of Approved Researchers and to publishing the findings of their research (including on the ONS Approved Researcher pages). ONS will work with RSA and Highways England to consider the development of a case study setting out further information on the research methodology and outcomes, and their impact on public service delivery and policy making.

B7 How will the findings of the research be disseminated?

The results will be published by Highways England and a link to these will also be included on the ONS Approved researcher web pages in accordance with the Approved Researcher transparency criteria.

B8 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.

None

Section C Details of Data Subjects

C1	Data subjects to be studied	
the	s the Study include all subsections of population (i.e. all ages, sex, ethnic ups etc	☐ Yes ☑ No If no please detail which subsections with justification(s) below
Subs	sections of the population (including vulneral	ple groups) the project focuses on: the research

9

NRS

focuses on those people that have committed road-related suicide. All ages, gender, ethnic groups will be included although it is likely that there will be a greater focus on people of age where they can lawfully drive.
Justification for focusing on these subsections or groups:

Please detail consent given to use data specified in section B2

No explicit consent gained so the ONS Approved Researcher gateway is being used.

If you are using data held by a third party please detail how you will obtain this

N/A

UK Statistics Authority

National Statistician's Data Ethics Advisory Committee

NSDEC(17)05

Web-scraping guidance for ONS staff

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

Ms Lan Benedikt

10 AOB

UK Statistics Authority

National Statistician's Data Ethics Advisory Committee

NSDEC(17)06

Administrative data based population and household estimates

This project is undergoing major revisions and will be published in due course