



The HMRC Datalab: New opportunities for the research community to use government administrative and survey data

Daniele Bega and Lucy Nicholson
HMRC Datalab
PEUK Conference
18th June 2013

Aims of the Presentation

- Overview of HMRC, the information we hold and how it can be accessed
- Describe the opportunities the Datalab offers for academic researchers
- The HMRC Datalab: details on governance, data available and how to apply
- Background on the legal context in which the Datalab and HMRC operate
- Tax Administration Research Centre

1) HMRC Data Landscape

What is HMRC?

- UK's tax authority
- Non-ministerial Department
- Established by Act of Parliament in 2005
- Reporting to Parliament through our Treasury Minister who oversees our spending
- HMRC/HMT partnership

HMRC's role

We **collect** and **administer**:

Direct taxes – paid by you or your business on money you earn or capital you gain.

- Capital Gains Tax
- Corporation Tax
- Income Tax
- Inheritance Tax
- National Insurance Contributions

Indirect taxes - paid by you or your business on money spent on goods or services.

- Excise duties
- Insurance Premium Tax
- Petroleum Revenue Tax
- Stamp Duty
- Stamp Duty Land Tax
- Stamp Duty Reserve Tax
- VAT

HMRC's role (continued)

We **pay** and administer:

- Child Benefit
- Child Trust Fund
- Tax Credit

We **enforce** and administer:

- Environmental taxes
- National Minimum Wage enforcement
- Recovery of student loans
- Customs duties including collection of data on imports and exports

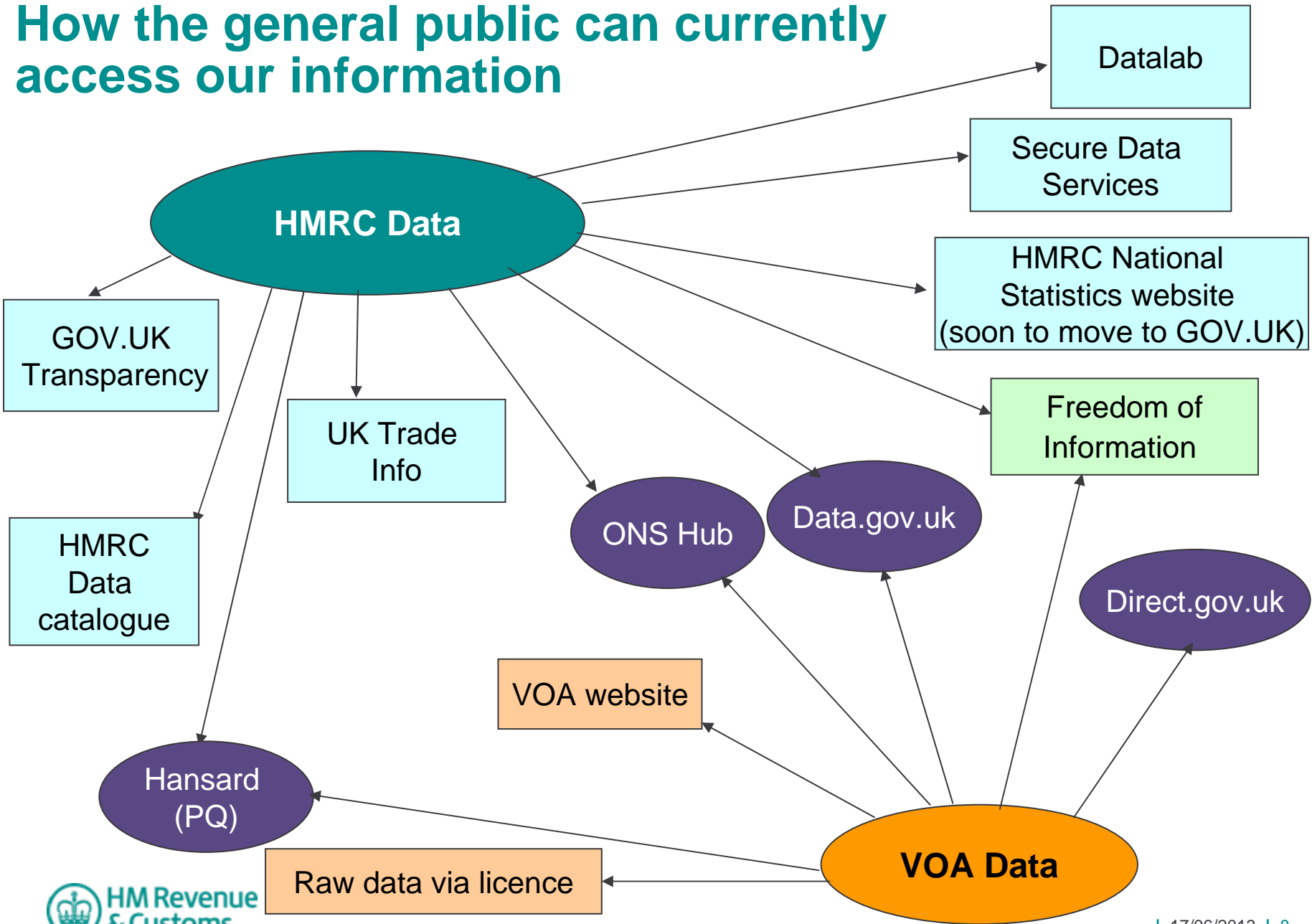
We are the parent department to the Valuation Office Agency (VOA) and its commercial arm – the District Valuer Services (DVS).

We also work with the Adjudicator's Office.

HMRC Data Landscape

Income Tax	Personal incomes	Property Transactions	Trade Statistics	Research and Development Tax Credits
Benefits in Kind	Charities	Non-domestic rating	Climate Change Levy	Child Benefit
Employee share schemes	Corporation Tax	Value Added Tax	Betting, Gaming and Lottery duties	Personal Tax Credits
Personal Pensions	Inheritance Tax	Insurance Premium Tax	Alcohol duties	Trusts
EIS & VCT	Personal Wealth	Air Passenger Duty	Hydrocarbon Oil duties	Tax Gaps
PEP, TESSA & ISA	Capital Gains Tax	Landfill Tax	Tobacco duties	Charitable Donations
Child Trust Fund	Stamp Duties	Aggregates Levy	Oil and gas production	National Insurance

How the general public can currently access our information



2) The HMRC Datalab

What is the Datalab?

The idea

A **secure** environment where researchers can access, free of charge, **confidential** taxpayer and customs **anonymised** data to undertake research that serves one of HMRC's functions and benefits the wider research community

The benefits

Improved evidence base and transparency with appropriate safeguards

How does the Datalab work?

Following the principles of the ONS Virtual Microdata Laboratory (VML) and Secure Data Services (SDS):

- Valid research purpose of interest to HMRC - **Safe projects**
- Trusted researchers from trusted institutions - **Safe people**
- Anonymisation of data - **Safe data**
- Controlled IT environment - **Safe settings**
- Disclosure controls of any output - **Safe outputs**

How the Datalab compares to other initiatives

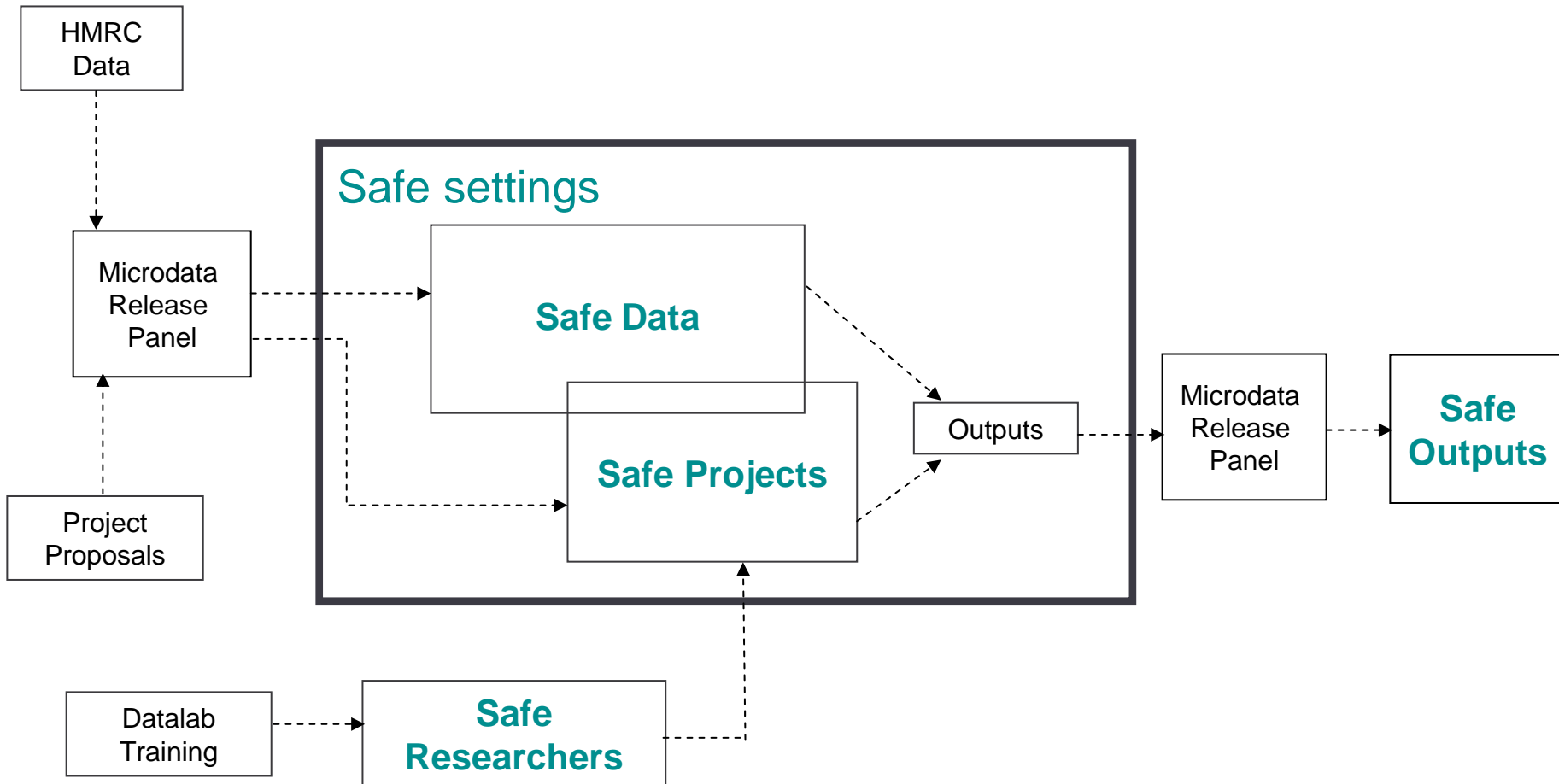
ONS Virtual Microdata Laboratory (VML)

- Information on VML server available on multiple sites
- VML granting access to Other Government Departments and non academic
- VML charges for non academic use

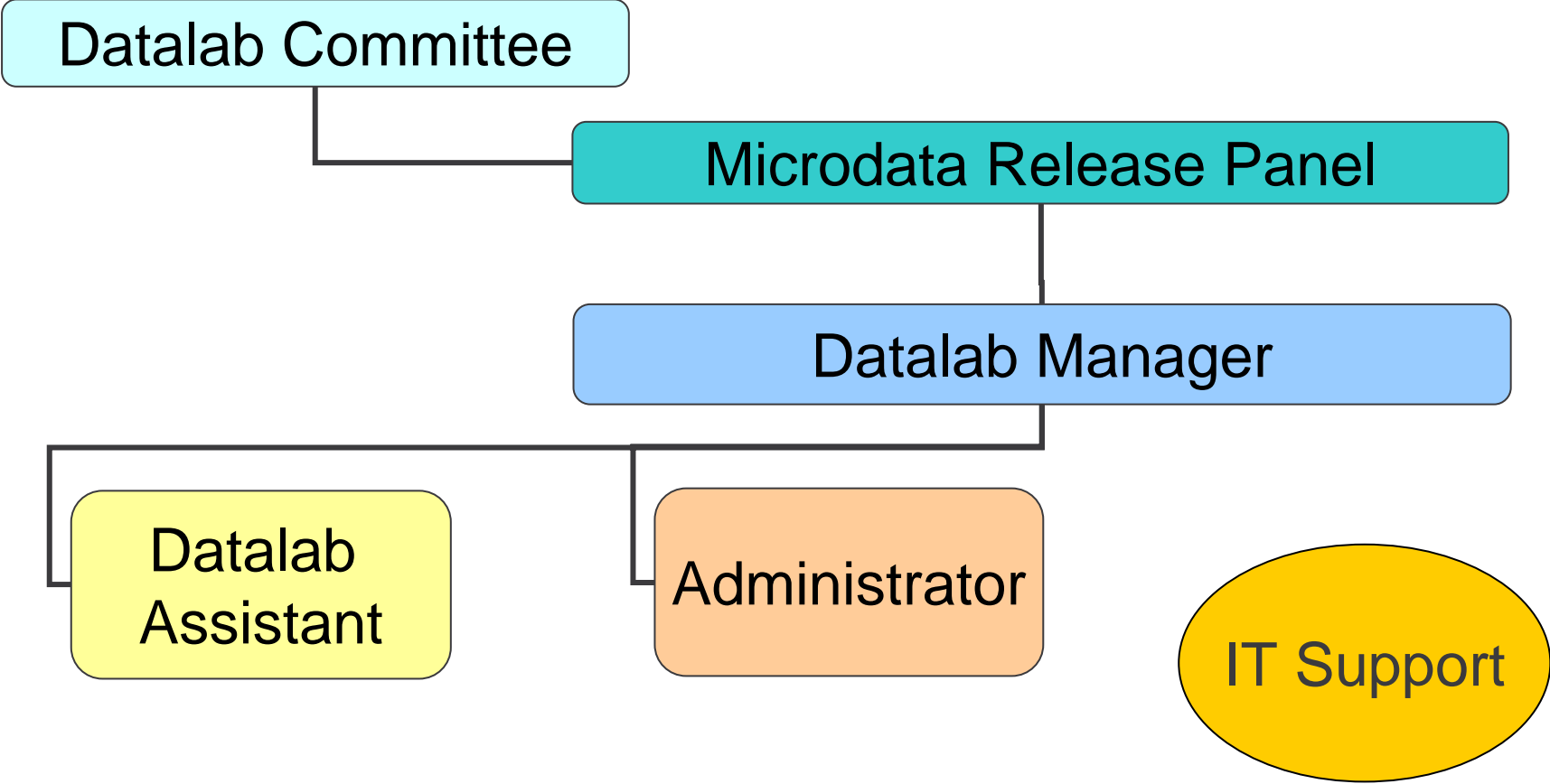
Secure Data Service (SDS)

- Researchers analyse the data remotely from their institutional desktop or in a safe room

Our Model



Governance of the Datalab



Data available

Current datasets:

Corporation Tax	Value Added Tax
Stamp Duty Land Tax	Trade Statistics
HMRC Customer Survey	Tax Credits
Survey of Personal Incomes (Public Use Tapes)	Partnerships
PAYE data	Self Assessment
	Compliance Perceptions Survey

We have uploaded FAME in the Datalab, but access is only allowed to institutions who already have a licence and subject to the provider's (BvD) authorisation

- Various lookup tables to join datasets
- HMRC Data Catalogue available on our website

The Datalab environment

- 8 computers - (64-bit PCs, 2 with 24GB and 6 with 32GB RAM)
- Multi Terabyte Server Capacity
- Software including:
 - StataMP 12
 - SAS 9.3
 - PASW Modeler 13 (Clementine)
 - MS SQL Server 2005 & 2008
 - MS Office 2003 and 2007 inc. Access, Excel, Word & Powerpoint
 - Data Matching tools e.g. DQ Global (Match)

Some institutions who have applied to use the Datalab

- Oxford University
- London School of Economics
- Nottingham University
- Imperial College
- Warwick University
- IFS
- Essex University

With 18 projects approved so far

Producing papers including:

- Corporation Tax in the United Kingdom (Michael Devereux & Simon Loretz)
- The Elasticity of Corporate Taxable Income: New Evidence from UK Tax Records
(Michael Devereux, Li Liu & Simon Loretz)
- Housing Market Responses to Transaction Taxes: Evidence from Notches and Stimulus in the UK
(Michael Best & Henrik Kleven)

The Datalab and the Legislation

- The legislation **restricts the use academic researchers can make of datasets in the Datalab**
- The Commissioners for Revenue and Customs Act (CRCA) is the Act of Parliament that created HM Revenue and Customs in April 2005. The legislation applies to **all HMRC officers**, whether they were former Inland Revenue or former Customs and Excise staff.
- Section 18 of CRCA makes clear that we **must not give** ('disclose') HMRC information to anyone, unless we have lawful authority to do so. This includes Other Government Departments and their agencies, local authorities, the police or any other public bodies.

Exemptions

- For the purposes of HMRC's functions

HMRC Functions – Datalab Interpretation

We allow research proposals that fall into one of these categories:

Compliance

Helping to encourage taxpayers to fulfil their requirement to comply with the law

Understanding Customers

Increased HMRC focus on understanding our customers' behaviour

Operational Efficiency

Helping to improve how HMRC operates and conducts its work

Policy Development

Support policy development, review effectiveness, learn lessons and improve policy making

Other Issues

Anything else that might be deemed relevant to HMRC's overall functioning and might build capability

HMRC Research Agenda

- Administration/ efficiency improvements
- Administrative burdens/ customer-related research
- Compliance
- Data quality issues and survey design
- Tax policy

Further details available on our website:

<http://www.hmrc.gov.uk/datalab/research-topics.htm>

The Datalab in practice?

Researchers submit their application to the Datalab inbox

The Microdata Release Panel assesses and scores the proposal

Often, researchers are invited in to discuss their proposal with HMRC analysts and policy colleagues

Final decision is made by the Datalab Committee

Proposal assessment process

- Once assessed by the MRP, proposals are reviewed on a quarterly basis by the Datalab Committee. The Committee grants final approval for all projects.
- In order for proposals to be considered in a timely manner, the following deadlines apply:

Datalab Committee Meeting	Deadline for project proposal submissions for this meeting
12 July 2013	14 June 2013
11 October 2013	13 September 2013
11 January 2014	13 December 2013

The Datalab in practice?

If approved, contracts are exchanged, data is prepared and date to commence research is decided upon



Datalab Team provide training course (on Datalab and statistical disclosure control)



Researchers are free to begin their project



All outputs are assessed by HMRC staff for potential disclosure issues before being released

How do I apply?

- **Contact the Datalab team:**
hmrc.datalab@hmrc.gsi.gov.uk
requesting for an application form
- **Submit a proposal**
- **Full details at**
<http://www.hmrc.gov.uk/datalab/apply.htm>

The Application Form

Personal details and contact information of lead researcher

Summary of proposed research

Description of the project and aims, as well planned duration of access

Research design and methods

Outputs and dissemination

Including publication plans and peer-review

Datasets

Data requirements, including details on whether any data matching needs

HMRC Benefit

How the project serves one of HMRC's functions and future engagement

Next steps

Consolidate the success of the Datalab:

- Continue to improve the experience of Datalab users
- Promote the Datalab amongst the research community
- Encourage discussion between HMRC and the research community
- Ensure successful collaboration with the Tax Administration Research Centre (TARC)
- Continue to collaborate with other Data Services
- Outline the role of the Datalab in the context of ADS and ADRCs
- Accreditation

3) The Tax Administration Research Centre

The Tax Administration Research Centre (TARC)

A jointly funded research centre (ESRC/HMT/HMRC), focusing on tax administration with the potential for:

- Building academic engagement and capacity, conducting longer-term research; enhancing the evidence base and enhancing in-house and academic skills;
- Providing high quality, policy relevant, multi-disciplinary research

Centre directed by the University of Exeter, in partnership with the IFS

Launched in January 2013, currently working on a research programme for the first year

Research around four areas: Analysis and Simulation, Estimation and Evaluation, Economic and Social Experimentation and Qualitative Analysis



Thank you

Datalab team: datalab@hmrc.gsi.gov.uk