



HM Revenue
& Customs

Sharing and Publishing Data for Public Benefit

Summary of Responses

10 December 2013

Contents

1	Introduction	3
2	Responses to Chapter 2 Proposals	8
3	Reponses to Chapter 3 & 4 Proposals	20
4	Next steps	34
Annex A	List of stakeholders consulted	35

On request this document can be produced in Welsh and alternative formats including large print, audio and Braille formats

1. Introduction

Background

- 1.1 Information held by the public sector is among the most useful and valuable to be found anywhere. This is why the UK Government is at the forefront in making a step change in the availability of data held by the public sector, with the intention to deliver significant public benefits.
- 1.2 In June 2013, the Government signalled its continued commitment to realising these benefits with two major steps. Firstly, the Government published its response to the Shakespeare Review of Public Sector Information, setting out a framework for pursuing this agenda in the public sector. Secondly, the UK helped secure the G8's Open Data Charter, which establishes the presumption that the data held by governments will be publicly available, unless there is good reason to withhold it.
- 1.3 Her Majesty's Revenue and Customs' (HMRC) relationship with businesses and individuals is unique. This is reflected in the scope and depth of the information HMRC collects, creates and protects on behalf of taxpayers in carrying out its departmental functions. HMRC recognises that it is important for the department to play a full part in the Government's Open Data agenda, using its information to improve transparency and promote economic growth. It also wishes to improve data sharing with other public sector bodies to deliver better services across the public sector.
- 1.4 HMRC was created by the Commissioners for Revenue and Customs Act 2005 (CRCA). This legislation provides strong protection for the information that HMRC holds. HMRC officials are prohibited from sharing information except in the limited circumstances set out in the CRCA. This legislation enshrines the core principle of what is often described as 'taxpayer confidentiality'. HMRC is committed to maintaining this important principle, which is essential to the effective operation of the tax system because it supports compliance and willing

cooperation. It is therefore paramount that any data release has appropriate safeguards.

- 1.5 The CRCA prohibition on disclosure applies to all of HMRC's information including non-identifying (general, aggregate or anonymised) information as well as information on identifiable individuals or legal entities. As a result, it is arguable that for non-identifying information the current disclosure restrictions afford more protection than is necessary.
- 1.6 General and aggregate information could be shared more readily and widely without compromising customers' confidentiality. This could help deliver more effective and efficient public services, for example, to help inform policy to improve social mobility.
- 1.7 In addition, wider policy outcomes could be delivered if the current legal restrictions were relaxed to allow anonymised information to be used for research and analysis purposes beyond HMRC's functions. This could be achieved while using appropriate safeguards to protect confidentiality similar to those already in place within the HMRC Datalab for existing research projects.
- 1.8 Finally, there is the potential to deliver wider public benefits through the release of specific data sets, which may currently be prohibited under HMRC's legislative framework. One such dataset is the register of all businesses registered for VAT (not including any financial or tax payment data). Such proposals raise further important questions, including how to ensure appropriate safeguards to protect confidentiality and privacy.
- 1.9 HMRC therefore carried out a consultation over the summer of 2013 on three specific proposals that it identified had the potential to deliver significant public benefits:
 - removing HMRC's legal constraints on sharing general and aggregated information, and relaxing HMRC's legal constraints on sharing

- releasing basic non-financial VAT registration data as open data; and
- sharing more detailed (but still non-financial) VAT registration data on a more controlled and restricted basis for specific purposes, such as credit referencing.

Details of the consultation

- 1.10 HMRC published the consultation document *Sharing and publishing data for public benefit* on 17 July 2013. The consultation was formulated in open terms so that it would be possible to explore fully all the issues to inform decisions. The consultation closed on 24 September 2013.
- 1.11 Chapter 1 of the consultation document set the context and the Government's plans to increase the availability of public data. HMRC already lawfully shares information widely across the public sector and publishes a wide range of non-identifying information. The chapter set out how this might be broadened and made it clear that primary legislation would be required to implement any of the proposals set out in the consultation document.
- 1.12 The consultation sought responses to specific proposals set out over the following three chapters. Chapter 2 concentrated on the sharing and publishing of general and aggregate data and the disclosure of anonymised individual level data under secure controlled conditions. Chapter 3 put forward a proposal for public release of a limited and selective extract of the VAT register, with options to reduce or increase the level of disclosure about VAT traders. Chapter 4 put forward proposals for a controlled release of a fuller VAT registration data set to qualifying parties for nominated purposes (e.g. anti-fraud checking), subject to rigorous controls.
- 1.13 The consultation document also noted the intention that any resulting legislation would be 'permissive' in nature. This means that HMRC would be able, but not obliged to share information. The importance of the core principle of taxpayer

confidentiality was underlined, along with the need to ensure appropriate safeguards were in place. This includes the criminal sanction as a strong safeguard against the unlawful disclosure of identifying information.

1.14 A research project is being carried out alongside the consultation to assess potential external uses of VAT registration data, possible benefits and issues arising, to help inform decisions on these particular proposals.

1.15 More details on the consultation can be found at: <https://www.gov.uk/government/consultations/sharing-and-publishing-data-for-public-benefit>

General responses

1.16 HMRC received 54 written responses from a range of respondents. Here is the breakdown by respondent type (this is indicative – some respondents could fit into more than one category):

Type of Respondent	Number of Responses
Individuals (including 6 academics)	16
Professional / representative bodies	12
Government and public bodies	10
Privacy / open data groups and users	7
Small and medium enterprises (SMEs)	3
Business information providers (e.g. credit reference agencies - CRAs)	3
Other large businesses	2
Research organisations	1

1.17 During the course of the consultation, HMRC held a number of meetings with interested parties (including, for example, various representative and

professional bodies) to discuss the proposals. Feedback from these meetings and subsequent written responses have been considered as part of the consultation exercise.

- 1.18 HMRC wishes to thank those who responded to the consultation document and recognises the time and effort that went into the comments and contributions, which have informed consideration of the policy options.

2. Responses to Chapter 2 Proposals

2.1 This chapter provides an overview of the responses received during the consultation to the proposals at chapter 2 of the consultation document and goes on to provide additional detail on the responses to each of the five questions asked.

Overview of comments

2.2 Many respondents noted that they welcomed the opportunity to comment on the detail of the proposals. In general, respondents were positive about the proposals to reduce legal constraints on sharing general, aggregate and anonymised information.

2.3 Chapter 2 of the consultation document presented two specific proposals:

- to remove HMRC's legal constraints on sharing general and aggregate information; and
- to relax HMRC's legal constraints on sharing anonymised individual level information with accredited third parties for the purposes of research and analysis.

2.4 A significant majority of respondents were either in strong agreement or agreed in principle to both of the proposals. All respondent sectors either agreed strongly or agreed in principle with the proposal to remove legal constraints on sharing general and aggregate data while qualifying their responses with a focus on the need to mitigate risks of disclosure that are broadly in line with the safeguards proposed in the consultation document. With regard to the second proposal, to relax legal constraints on sharing anonymised individual level data, most sectors were again strongly in favour. However, a greater degree of caution was expressed on safeguards than was the case for general and aggregated information. Again, the suggestions offered were broadly in line with the enhanced safeguards HMRC is proposing in relation to anonymised data.

2.5 Some interesting suggestions were received on increasing transparency of research and analysis work, which HMRC will be considering further.

Responses to chapter 2 questions

Question 1: Do you agree that the legal constraints on sharing general and aggregate information should be removed, on a permissive basis, in respect of:

(i) HMRC making information generally available through publishing?

(ii) HMRC sharing information with specific third parties to deliver public benefits wider than HMRC's functions? Please give reasons for your answers.

There were 30 responses to this question

2.6 There was a strong consensus that legal restrictions on publishing and sharing general and aggregate data should be removed to allow wider data-sharing on a permissive basis. This was subject to having appropriate safeguards in place to protect taxpayer confidentiality. Some respondents commented that the data should be freely available and possibly published under the Open Government Licence.

2.7 CRAs, government and public bodies, and privacy / open data groups saw the most benefits in removing the legal constraints. These included:

- a reduction in fraud;
- helping the finance sector to establish who might be at risk from money laundering;
- increasing HMRC transparency and therefore trust in the Department;
- private sector organisations could use the data to inform their decision-making to help build new businesses and better services;
- supporting and improving policy making; and
- helping to avoid duplication of data collection across government.

2.8 On a more cautious note, a minority of respondents suggested:

- they were not convinced that sharing data with specific third parties would deliver significant public benefits; and
- a need for greater clarity on the interpretation of “public benefit”, so that data sharing does not benefit only “public bodies”.

2.9 Many respondents commented on the importance of maintaining privacy and emphasised the need to ensure that data remains anonymised and that individuals cannot be identified.

2.10 Some respondents had concerns that there would not be a level playing field if some groups had privileged access to HMRC data.

2.11 Some respondents offered further suggestions for safeguarding the confidentiality of taxpayers’ data including:

- closer working between HMRC and the Office for National Statistics (ONS) to determine the size of groups to be used in the data to protect identities;
- following the guidance in the Information Commissioner’s Office (ICO) anonymisation code of practice; and
- a ‘sunset clause’ in the legislation so that “after a suitable period appropriate consideration would be given to whether the relaxations had achieved their purpose, and whether any changes were meeting the twin objectives of driving public benefit whilst maintaining taxpayer confidentiality.”

Government response to question 1

The consultation proposed removing the current legal restrictions on sharing general and aggregate information. It is intended that any disclosure would be on a permissive rather than a mandatory basis. This would also enable HMRC to manage its costs, ensure its core function is not compromised, and allow robust prioritisation of disclosures by direct linkage to public benefit.

Responses to this question are generally positive. However, responses also highlight understandable concerns about the need to safeguard privacy and confidentiality of information.

The Government has decided to proceed with the proposal to remove legal restrictions that currently limit HMRC's ability to share general and aggregate information for public benefit.

This move will also address the current anomaly whereby HMRC may be legally required to provide aggregate information requested by a citizen under the Freedom of Information Act (FOIA) but has to deny a request for the same information from another government department in instances where CRCA does not provide for the disclosure.

HMRC accepts that it will need to be clear and transparent on what is meant by 'public benefit'.

In addition, HMRC accepts that it will need to set in place comprehensive governance, policies and processes, including the evaluation of benefits, risks and costs of a disclosure, before any data is shared or published. This will include ensuring that any public benefit gained from publishing the information will outweigh any benefit from withholding the information.

Question 2: Do you agree with the proposed safeguards on the proposal to share general and aggregate data? Should any further controls be considered on what can be shared, with whom or how?

There were 26 responses to this question

2.12 Most of the respondents agreed or agreed in principle with the proposed safeguards on general and aggregate data.

- 2.13 Only two respondents disagreed, expressing strong concerns that sophisticated modelling techniques could be used to identify individuals. They felt that HMRC needed to convince them that it would review the data against the newest modelling techniques before release.
- 2.14 A large number of the respondents and professional/representative groups in particular, highlighted the importance of robust safeguards, suggesting other safeguarding techniques and organisations as possible exemplars. In particular:
- it was suggested that HMRC should adopt the protocols followed by the ONS and the Economic and Social Research Council (ESRC) and those recommended in the Administrative Data Taskforce report;
 - HMRC was pointed to the Organisation for Economic Cooperation and Development's report 'Promoting Access to Public Research Data for Scientific, Economic and Social Development'; and
 - the VAT Information Exchange System (VIES) was mentioned by a few respondents as an example of taxpayer data being available but protected, since people have to know certain information already to access the data.
- 2.15 One respondent commented on the importance of considering the value of the criminal sanction and questioned whether it acted as a sufficiently strong deterrent.
- 2.16 Others suggested further safeguards, including:
- ensuring a senior responsible officer is employed to check the data before it is published to make sure no individuals can be identified;
 - allowing taxpayers to claim damages from third parties (or HMRC where that is not possible) if their information is compromised;
 - data should be of a suitably high level to minimise the possibility of individuals being identified; and

- aggregate data should be regularly reviewed to ensure it cannot be used for identifying specific taxpayers.

2.17 A few respondents also suggested that the process used by HMRC to decide how and when to publish data should be open to challenge and audit.

Government response to question 2

HMRC is committed to maintaining the important principle of ‘taxpayer confidentiality’, which is essential to the effective operation of the tax system because it supports compliance and willing cooperation. It is therefore paramount that any data release has appropriate safeguards.

The consultation proposed a number of safeguards to prevent customers’ identities being deduced from aggregate data and to protect delivery of HMRC’s core business objectives.

Responses to this question highlight the importance that many respondents attach to robust safeguards. Most respondents agreed with the safeguards proposed in the consultation and some suggested additional controls.

HMRC already publishes a wide range of general and aggregate information and ensures that safeguards keep in step with industry best practice. Statistical disclosure tests are applied with the aim of ensuring that it is not possible to deduce information about identifiable persons from aggregate information.

HMRC has already put into practice the safeguards set out in the consultation and will want to ensure that they continue to apply to any data release.

HMRC acknowledges the need to work closely with recognised experts, such as the ICO and the ONS, to design and install safeguards that continue to reflect “best practice”.

HMRC will also adopt additional safeguards suggested by respondents where commensurate with the level of risk.

To aid transparency and maintain confidence in safeguards applied, HMRC intends to publish information on data sharing governance, policies and processes.

Question 3: Do you agree that HMRC should be able to share anonymised individual level data for the purposes of research and analysis to deliver public benefits wider than HMRC's own functions? Please give reasons for your answer.

There were 28 responses to this question

- 2.18 With the exception of one individual, all respondents agreed or agreed in principle that anonymised individual level data should be made available for research and analysis in the public benefit.
- 2.19 A large number of respondents from across all sectors detailed benefits they expected to result from sharing anonymised data, including improving policy development and research and cutting down on the costs of large surveys. The benefits to policy making were particularly noted by government and public bodies.
- 2.20 A few respondents suggested that the data could be used for wider issues such as assisting with social mobility and addressing poverty.
- 2.21 On a more local level, some respondents were clear that the proposals would help local government planning and devolved administrations.
- 2.22 However a concern was expressed that as individual taxpayers could be more easily identified from this type of information, it should only be shared with other government departments and not with private companies.

2.23 A few respondents suggested that decisions to publish data should be made public. These respondents also felt that the results of any research undertaken by the third parties should be shared publicly and that the HMRC Datalab should be open to any party willing to fund the work needed and to publish its results.

2.24 Finally, a few respondents suggested going beyond the proposals, in order to enable government departments to share identifying data.

Government response to question 3

The consultation proposed removing the current legal restrictions on sharing anonymised individual level data for the purposes of research and analysis to deliver public benefits wider than HMRC's own functions.

Responses to this question are broadly positive. However, a significant number of responses highlight concerns around the dangers of individual identities being disclosed inadvertently.

HMRC recognises that stronger safeguards need to be built in to this proposal to protect taxpayer confidentiality. This will be considered further in the response to question 4.

HMRC already makes anonymised individual level customer data available via its 'Datalab' to accredited researchers for the purposes of research where there is a direct benefit to HMRC's functions.

However HMRC is unable to realise the potential to deliver public benefits that fulfil wider policy objectives beyond its own functions.

The Government has decided to proceed with the proposal to remove the legal restrictions that currently limit HMRC's ability to share anonymised individual level data for the purpose of research and analysis and deliver public benefits wider than HMRC's own functions. But they accept that this must be done only where there are sufficient safeguards in place to protect taxpayer confidentiality.

Question 4: Do you agree with the proposed safeguards on the proposal to share anonymised individual level data? Should any further controls be considered on what can be shared, with whom or how?

There were 27 responses to this question

- 2.25 Almost all respondents agreed or agreed in principle to the proposed safeguards and a number suggest further controls. A small number of respondents suggested that the proposed safeguards provided more protection than in fact would be needed. Only one respondent was entirely against the concept of sharing anonymised data, due to concerns that individuals would be identified.
- 2.26 A number of respondents suggested that HMRC should follow good practice approaches adopted by other organisations to sharing anonymised data. Some of the organisations mentioned were:
- the ONS and its Virtual Microdata Laboratory;
 - the ESRC's principles on making individual level data available to researchers; and
 - the recommendations in the response to the Administrative Data Taskforce.
- 2.27 Respondents noted that these organisations could help with the linkage of datasets as well as protecting privacy and confidentiality.

- 2.28 Some respondents thought that the information should only be shared with specific qualifying parties, including accredited researchers and other government departments. It was suggested that there should be a tiered approach to releasing data, with different levels of data available to different parties. A few respondents proposed an accreditation or vetting process, with one respondent wanting a list of accredited users to be published.
- 2.29 Some respondents sought greater transparency with respect to HMRC Datalab arrangements and specifically in relation to how data would be protected once it had left the Datalab and the types of restrictions that would be put in place. It was also suggested that any research undertaken in the Datalab should be published.

Government response to question 4

HMRC is committed to the core principle of taxpayer confidentiality and it is vital that any release of anonymised individual level data has appropriate safeguards in place to allay concerns around the dangers of individual identities being disclosed inadvertently.

The consultation proposed a number of safeguards and responses to this question have shown that respondents take this issue seriously. Most respondents agreed with the suggested safeguards and some suggested additional safeguards to those mentioned in the consultation and organisations as possible exemplars.

HMRC currently operates its Datalab under strictly controlled and secure conditions and in line with other UK accredited research data centres. Any data sharing enabled as a result of reducing legal constraints on sharing anonymised individual level data would follow the best practice principles already in place within the Datalab.

HMRC is looking closely at the additional safeguards and options suggested by respondents and will adopt them if they are commensurate with the assessed level of risk.

As mentioned in the Government response to Question 2, HMRC intends to publish information on data sharing governance, policies and processes.

Question 5: How should the generation and release of anonymised or aggregated data be funded? Please give reasons for your answer.

There were 26 responses to this question

- 2.30 A large proportion of the respondents were of the view that this data should be made available without charge and that any costs related to publishing or sharing data should be borne by HMRC, particularly where data is provided to other government departments. In most of the responses that favoured free data provision, there was a caveat that the data should only be free providing it is of benefit to the public. A few responses were in favour of using the Open Government Licence to release data free of charge.
- 2.31 A number of respondents suggested that a charge should be levied to fund publishing or sharing data in circumstances where a data request takes up significant HMRC resource. Most did not suggest a specific amount to be charged, although one respondent mentioned the idea of adopting the same charging process as Companies House. Where respondents did suggest charging, most made it clear that the charge should not be so large that it prevented some organisations, particularly small businesses, from being able to access the data.
- 2.32 Some respondents felt that the party requesting the data should provide any necessary funding, and this should include other government departments.

2.33 Other suggested sources of funding for the generation and release of aggregate or anonymised data included research councils, higher education councils and charities as well as government. It was also suggested that Administrative Data Research Centres could provide a suitable vehicle for HMRC data publication and sharing.

Government response to question 5

HMRC is reviewing the ideas raised by respondents for funding the proposed data sharing options. It is clear that there is no consensus for one blanket approach to charging, which suggests that a more granular approach depending upon purpose, requestor and data type/volume should be explored. HMRC will consider charging options as it develops its data-sharing policy/strategy and considers the likely resources needed for data sharing requests.

3. Responses to Chapter 3 & 4 Proposals

Consultation proposals

- 3.1 Chapters 3 and 4 of the consultation *Sharing and publishing data for public benefit* sought comment on proposals for the release of some VAT registration data, i.e. data identifying businesses registered for VAT, but specifically excluding disclosure of any financial or tax payment data.
- 3.2 Chapter 3 asked for comments on a core proposal to release publicly a selective extract of VAT registration data (VAT registration number (VRN), trading name and Standard Industry Code (SIC) classification number). This was termed Option A. Comments were also sought on proposals to permit business owners to opt out of public disclosure (Option B) and/or opt in to disclosure of additional information such as contact details (Option C).
- 3.3 Chapter 4 asked for comments on a proposal for controlled release of a fuller set of identifying data provided by VAT registered businesses (set out at Annex C of the consultation document), within a rigorous framework of safeguards. Data would only be made available to qualifying parties for specific purposes – examples given in the consultation document were credit scoring and anti-fraud checking – and the questions sought additional views on suitable purposes.

Overview of responses

- 3.4 A total of 43 respondents commented on the release of VAT registration data. There was majority support for the principle that public authorities should be able to use information they hold for the wider public benefit, provided any disclosure done to achieve this is handled responsibly and securely. There was also support for the observation by one respondent that business activity generally takes place in a more public arena than private life and that making a simple distinction between ‘personal’ and ‘non-personal’ information is not always straightforward.

- 3.5 However, the principal characteristic of the responses was the diversity of the views expressed, and in some cases their incompatibility. Respondents did not share a common view on the criteria by which the Government should evaluate VAT registration data release proposals, nor on the impact of the proposals under different criteria (e.g. the net benefits which might be generated by a data release).
- 3.6 In broad terms, the responses fell into four groups.
- 3.7 The first group supported the release of VAT registration data because of the benefits they believed would flow from this, in particular improved access to and availability of credit for SMEs. While often sympathetic to an open data release, they generally believed that a controlled data release would maximise net economic benefits and minimise risks, with some suggesting it might also be a precursor to a subsequent open data release. This group largely consisted of businesses and their representative groups (including credit providers, business information providers and SMEs).
- 3.8 The second group supported an open data release of VAT registration data (in some cases preferring the publication of different data fields). Most in this group were opposed on principle to a controlled data release. Some felt that an open data release would generate more benefits than a controlled release.
- 3.9 These respondents felt that the key issue was what data could appropriately be released by the Government, and that the primary criterion for this decision should be the degree of safeguards afforded to personal privacy. Hence, more data should be released on companies (which had limited expectations of privacy) and less on individuals. They believed that once this criterion had been applied, the releasable data should be made available to all as open data. Businesses should have the option of volunteering the release of additional data on top of this. This group was mainly composed of civil society interests, including open data groups, and potential commercial users of open data.

- 3.10 The third group had reservations about any data release, either controlled or open. Their reasons were not uniform, and included concern that economic benefits might not be achieved; that release of registration data might be the ‘thin end of the wedge’, subsequently leading to further data releases, including of taxpayer financial data; and a view that data, including in some cases personal data, provided under compulsion for tax administration, should not be released on principle, and that to do so might weaken tax compliance. This group was drawn principally from professional bodies representing tax and legal advisers.
- 3.11 Responses from government departments, devolved administrations and other public bodies made up a fourth group. These generally supported release by whichever means would make the data available for public usage, while applying appropriate safeguards, such as those discussed in chapter 2 of this document.
- 3.12 Responses to chapters 3 and 4 of the consultation are explored further below, by reference to each question. The diversity of responses raises important issues. These require careful consideration to ensure a robust outcome is achieved, based on clear and appropriate criteria. Accordingly, this chapter is limited to summarising responses received and the issues raised. The Government’s interim response is set out at the end of this chapter.

Responses to chapter 3 questions

- 3.13 Questions 6 to 10 were in respect of the proposals to publish limited VAT registration data (i.e. to release it to the public as open data).

Question 6: What potential uses and benefits do you think would arise from publishing VAT registration data?

There were 30 responses to this question

3.14 A majority of respondents believed that there would be benefits from publishing VAT registration data, giving as examples those cited in the consultation document and identifying additional benefits. Few respondents attempted to quantify benefits, although most believed that the level of benefit would increase with the amount of data made public.

3.15 *(a) Impact on data for credit scoring and access to credit*

Respondents – including CRAs, finance and business representatives – said that the availability of VAT registration data would provide better and more accurate identifying information which, when combined with other business data, would improve assessment of the creditworthiness of VAT registered businesses, thereby improving access to and volume of credit, in particular trade credit. These respondents said that VAT registration data would provide a valuable data source on unincorporated and recently incorporated businesses, where there is a dearth of information. The absence of reliable information, rather than the presence of any negative indicators of business viability, is a key factor in lowering credit scores. One response stated that credit scores for unincorporated businesses could be an average 40% lower owing to paucity of data. It was also said that banks often rely on information from CRAs in considering borrowing requests by SMEs, and that it can be difficult to match CRA data where these businesses are unincorporated or recently incorporated.

3.16 A smaller number of respondents were sceptical that there was evidence to support the argument that VAT registration data would materially improve credit scoring and access to credit, or were unconvinced by the case set out in the consultation document.

3.17 *(b) Reducing fraud*

A number of respondents said that increased data on VAT registered businesses should improve the accuracy of checks to combat fraud and money laundering. One noted that fraudsters often exploit asymmetries in information availability. Greater openness could provide a remedy to combat data 'black markets'. However, some others argued that more evidence for benefits in

relation to fraud checks was needed and/or suggested that benefits might be outweighed by an increased risk of identity fraud.

3.18 (c) *Other benefits*

Consultation respondents also identified the following additional potential benefits from publishing some VAT registration data. Some are contingent on release of specific data fields (e.g. postcode):

- improving customer confidence in small businesses, by helping to provide assurance for suppliers and customers. VAT registration data could be an additional data source for services that allow consumers to check and review small businesses, such as tradesmen;
- providing more systematic availability of information about businesses e.g. by geography and industry sector. This would inform analysis within and outside government, including policy development, impact assessment and measurement of business support and regeneration measures, as well as of wider social, health and education policies. One response mentioned that such data might help combat local rating or benefit fraud;
- helping to identify businesses eligible for targeted support from central, local or devolved government, broadening access to such support; and
- facilitating new, innovative and as yet-undiscovered uses of VAT registration data.

Question 7: Do you believe that the proposed safeguards are sufficient to mitigate any risks arising from publication? What, if any, additional impacts may arise and what further safeguards should be considered?

There were 30 responses to this question

3.19 Respondents who supported an open release of VAT registration data felt that individuals' privacy could be appropriately safeguarded by limiting the data fields to be published and by permitting individuals to opt out of or into additional disclosures (see responses to questions 9 and 10 below). They also believed that publication would help mitigate impersonation fraud risks because

a reliable database of VAT registered businesses would improve detection of attempted fraud. They drew attention to experience in other European countries such as Slovenia, where the equivalent of the VRN is published as part of a comprehensive business registration database, and to Australia, where Business Number information is freely available. Several also pointed out that the name and address of any VAT registered business can be obtained for anti-fraud checking purposes, if one has a valid VRN for that business, by using the online VIES system established under the framework of EU rules for VAT (some suggested using VIES as a model for data release).

- 3.20 Respondents who opposed any publication of VAT registration data were not convinced by the discussion of privacy safeguards in the consultation document, taking the view that consent should be required. Concerns were raised about the consistency of any release of VAT registration data with both Data Protection and Human Rights Acts. Some respondents thought that public availability of VAT registration data might make commercial fraud easier.

Question 8: Do you agree with the core proposal (Option A) to publish the VAT registration number, trading name and industry classification? Please give reasons for your answer.

There were 34 responses to this question

- 3.21 A majority of respondents were prepared to support publication of a limited VAT registration data set. However, views differed as to the circumstances in which publication should occur and the extent of the data which should be published.
- 3.22 Some of those respondents whose preference was for a controlled VAT registration data release were also sympathetic to a open data release, albeit perhaps at a later stage and with less data released. It was seen as a complement to a controlled release.

- 3.23 Those respondents whose preference was for an open VAT registration data release mostly argued that there should be no controlled data release. Their reasons are given in the discussion of responses to question 11.
- 3.24 Respondents with a preference for an open data release had different views on the extent of the data which should be published. Some thought there would be little or no risk in releasing data such as address or postcode which might identify natural persons, while greatly increasing the benefits of use.
- 3.25 A majority, however, thought that it should not be possible to identify individuals from publication and personal data should only be released with those persons' consent. Accordingly, minimal data fields for natural persons (some of whom run their business from home and/or using their own name and personal contact details) should be published on a mandatory basis. Some suggested that no data should be published in respect of unincorporated businesses; others that only a list of valid VRNs should be published to facilitate verification of other business data. One response suggested that a phased approach might be possible, publishing corporate data first and then publishing the data of other traders after they had been notified.
- 3.26 Most respondents with a preference for an open data release sought the publication of additional data fields for companies (including, in various combinations, legal and trading names, incorporation number, address, contact details, and dates of registration and deregistration). The reasons were three-fold. First, that the utility of the published data set would be greatly increased by additional fields, for example to identify businesses eligible for support schemes, and it would increase the reliability of data matching with other sources (e.g. Companies House). Secondly, that businesses, or at least companies, did not have the same expectation of privacy as natural persons. Thirdly, that much of this additional data was already publicly available through Companies House, business directories, the VIES system mentioned under question 7, and from the businesses themselves (e.g. from invoices, websites and advertising).

3.27 Respondents who opposed any publication of VAT registration data were concerned that the suggested economic benefits might not be achieved (or could be outweighed by costs arising from increased fraud, increased nuisance, such as cold calling, or liability costs for data error or misuse). There was a belief that taxpayer data provided under compulsion for tax administration should not on principle be used for another purpose without consent and that doing so might weaken tax compliance. In addition, they had concerns that wider use of VAT registration data might be the ‘thin end of the wedge’, leading to more extensive data releases.

Question 9: What are your views on the proposal to allow exceptions or opt-outs to publication (Option B)?

Question 10: What are your views on the proposal to publish additional data fields on a voluntary basis (Option C)?

There were 29 responses to question 9, and 27 responses to question 10

3.28 The great majority of respondents, whether commercial, government, academic, or other, believed that introducing opt-in or opt-out elements to an open data release would be likely to increase complexity, increase costs for HMRC, and reduce the validity and utility of the data set by making it less consistent and comprehensive.

3.29 Nonetheless, a number of respondents supported a limited opt-out to any open data release, as a means of managing privacy concerns. Several respondents suggested that the impact of an opt-out could be minimised by adopting narrow criteria – for example restricting it to unincorporated businesses or to those which use their personal address for VAT registration. Another suggested allowing non-commercial entities (such as parish councils) to opt out. Several respondents said that a full list of VRNs should always be published. Only two respondents thought that companies might be permitted to opt out and then only in respect of directors’ contact details (reflecting similar provisions at Companies House).

- 3.30 Some respondents with a preference for an open data release preferred an opt-in approach as a way of securing consent from natural persons to the publication of their VAT registration data. However, a larger number felt that this would result in a non-random (because self-selected) and incomplete data set.

Responses to chapter 4 questions

- 3.31 Questions 11 to 15 were in respect of the proposals for a controlled release of VAT registration data.

Question 11: Do you agree with the proposal for controlled release of the full set of VAT registration data to qualifying parties? Please give reasons for your answer.

There were 34 responses to this question

- 3.32 Responses to this proposal were divided, reflecting the division in the responses to question 8.
- 3.33 Respondents with a preference for a controlled VAT registration data release, principally finance and business representatives, thought that this would maximise the net economic benefits by a significant margin. It would allow fuller registration data, covering both incorporated and unincorporated businesses, to be utilised, and to be applied to the most beneficial known uses (e.g. credit scoring). They also argued that a controlled release would protect privacy, as personal data (e.g. the names and addresses of sole traders) would only be available for the specified purpose. They thought it would help combat commercial fraud, as the data would assist in detecting discrepancies (e.g. in credit applications), and VAT fraud, as matching VAT registration data with commercial databases would help to identify businesses that had failed to register for VAT as required. One response suggested that controlled release for credit scoring purposes might provide an opportunity to encourage greater transparency by CRAs on scoring decisions.

- 3.34 Respondents who preferred an open data release rejected a controlled release. They believed that a controlled release of VAT registration data without consent would compromise privacy. They also thought that a controlled release would favour those entities with access to the controlled data over competitors and might be inconsistent with the principles of the G8 Open Data Charter. Some believed that a controlled data release would have a lower level of net benefits as it would support fewer uses and incur more release costs than an open data release.
- 3.35 Respondents who opposed any data release gave similar reasons for rejecting both controlled and open data release proposals (see responses to question 8). However, it was thought that a controlled release did not increase fraud risk, but did have the risk of data loss or misuse.

Question 12: What uses do you see for VAT registration data (see full list at Annex C) that would generate sufficient public benefit to justify controlled release? Please describe the uses and benefits in as much detail as possible.

There were 24 responses to this question

- 3.36 Respondents thought that the uses for VAT registration were broadly similar under both controlled and open data release proposals (see responses to question 6). It was recognised that a controlled release would probably result in a narrower range of uses, and that it might limit the development of innovative uses of VAT registration data. Respondents with a preference for a controlled data release thought that it would be more useful because the data set would be more comprehensive and larger than in an open data release.
- 3.37 Those opposed to any data release had concerns about the range of uses which might be permitted, for example direct marketing or tracing of persons who might not wish to be identified.

Question 13: What factors should be considered in determining the uses and parties that would qualify for release of the full VAT registration data set?

There were 23 responses to this question

3.38 Most respondents to this question had a preference for a controlled release. They thought that data should be released where public benefits (including wider economic benefits), could be identified and appropriate safeguards could be put in place. Controlled release should not be limited to CRAs, nor to the purpose of credit scoring. Data should be released to any qualified party (e.g. within government, to banks or other credit or business information providers). They noted that the parameters of permissible uses should be defined clearly in any new regulation and contractual relations between HMRC and each data user should specify data uses and safeguards.

Question 14: Do you believe that the proposed limits and safeguards are sufficient to mitigate any risks arising from controlled release? What, if any, additional impacts may arise and what further safeguards should be considered?

There were 22 responses to this question

3.39 As discussed in responses to other questions, respondents who had a preference for a controlled data release believed that strictly limiting the uses and users of VAT registration data would appropriately safeguard personal privacy, and that controlled use of the data would reduce overall fraud risk. Respondents who preferred on open data release disagreed with this approach to safeguarding privacy (arguing that individual consent should be required for any use of personal data). Some respondents who opposed any data release were concerned about the risk of personal data being lost or misused despite safeguards, suggesting that commercial entities may take calculated risks if there are opportunities to use data in additional ways such as marketing.

Question 15: How should a controlled data release be funded? Please give reasons for your answer.

There were 23 responses to this question

- 3.40 Respondents chose to offer comments on the funding of both open and controlled data release proposals under this question.
- 3.41 Most supporters of an open data release argued that data released by Government should be released free of charge under the Open Government License, in line with the Open Data Charter. However, a number of respondents felt that a charging regime might be appropriate. If so, it should be transparent, fair and should at most seek to reclaim costs. One of these responses observed that indirect costs such as software development should be recovered, to avoid a public subsidy of private benefits.
- 3.42 Respondents who supported a controlled release said that it would generally be appropriate for the Government to recover costs where data is shared with commercial entities.
- 3.43 Few respondents commented on the possibility that HMRC might seek other benefits such as enriched data in return for VAT registration data, although it was noted that such arrangements might exclude some potential users. One supporter drew an analogy with the arrangements for supply of data by some non-government suppliers to the Office of National Statistics (ONS). It was agreed that any such arrangements should be as transparent as possible.

Government response to questions 6 to 15

The Government and HMRC are grateful for the high level of responses to this consultation and the detailed consideration given to the proposals. The range and diversity of responses raise important issues. These require careful consideration to ensure a robust outcome is achieved, based on clear and appropriate criteria. Accordingly, this response to consultation has been limited to summarising responses received and issues raised. The Government will announce a decision after it has considered the issues fully.

Respondents to the consultation had differing views on the balance the Government should strike in choosing the criteria (including appropriate privacy safeguards) for evaluating VAT registration data release proposals. Respondents also offered a range of views on the net benefits likely to be achieved under different proposals and on the strength of evidence supporting them. HMRC is undertaking further work to assess the potential benefits and costs and will take this into consideration when preparing options and assessment criteria.

This includes the results of the current joint research project which is being undertaken with a group of CRAs to help assess potential external uses of VAT registration data, possible benefits, and issues which arise from such uses. Details of this research project were provided in Annex D of the consultation document. A further update on the project has been published on the consultation website at <https://www.gov.uk/government/consultations/sharing-and-publishing-data-for-public-benefit>.

Respondents had a range of views on the design of data release options (including the data fields which should be included in any open data release) and the appropriate safeguards. HMRC is considering these, the comments made on the use of opt-in and opt-out mechanisms to address consent issues and the cost-benefit implications of such mechanisms.

An announcement on next steps will be made in the new year. If the Government decides to release any VAT registration data, the intention would be to implement this through the same legislation which will be used to implement the release of aggregated and anonymised data. Draft legislation would be made available for scrutiny in the course of that process.

4. Next steps

- 4.1 This consultation generated many helpful, informative and constructive responses.
- 4.2 The Government has taken those responses on board, as well as the outcomes of meetings with stakeholders, to design its response as detailed in this document and the next steps it plans to take with regard to all the proposals suggested.
- 4.3 The proposals detailed in chapter 2 of the consultation (questions 1 to 5) are now at a stage where, thanks in part to the consultation responses, they can be taken forward by the Government.
- 4.4 The Government is now exploring legislative options for these specific proposals, with a view to publishing draft legislation in the new year. It is likely that the legislation would be introduced to Parliament in 2014/15 at the earliest.
- 4.5 The proposals in chapters 3 and 4 of the consultation document evoked a greater variety of responses and these need to be carefully considered, in the context of further research and work which is being undertaken to understand the impact of different options. It is important that the Government considers carefully the range of opinions on options for the release of VAT registration data and ways in which differences might be addressed. It is also important that the Government take into account the results of the current research project to assess potential external uses of VAT registration data, and the benefits, before making any decision on how to proceed. Further announcements will be made in early 2014 on the Government's plans in relation to VAT registration data.
- 4.6 If the Government decides to release any VAT registration data, the intention would be to implement this through the same legislation which will be used to implement the release of aggregated and anonymised data.

Annex A: List of stakeholders consulted

Applex Management
Association of Accounting Technicians
Association of Taxation Technicians
British Bankers' Association
Business Information Providers Association
Chartered Institute of Taxation
Creditsafe Group
Demographics User Group
Department for Business Innovation and Skills
Department for Work and Pensions
DueDil
Economic and Social Research Council
Experian
Federation of Small Businesses
Highlands and Islands Enterprise
Information Commissioner's Office
Inquiron
Institute of Chartered Accountants in England and Wales
Institute of Credit Management
Jordans Limited
Law Society
Local Public Data Panel, University of Southampton
Low Incomes Tax Reform Group
Moving Flow Limited
National Crime Agency (previously Serious Organised Crime Agency)
National Records of Scotland
Office for National Statistics
OpenCorporates/ Open Knowledge Foundation/ Open Rights Group (joint submission)
Open Data Institute
Open Data User Group
PricewaterhouseCoopers
Privacy International
Scottish Government
SSentif Intelligence
TaxPayers' Alliance
VAT Practitioners Group
Wellcome Trust
Welsh Government

16 individuals also responded, including 6 academics