



# UK public unhappy with the way their personal data is managed

Findings from a survey of public views on data management models

---

Helen Kennedy, Todd Hartman and Robin Steedman, University of Sheffield  
Rhianne Jones, BBC Research and Development

# EXECUTIVE SUMMARY

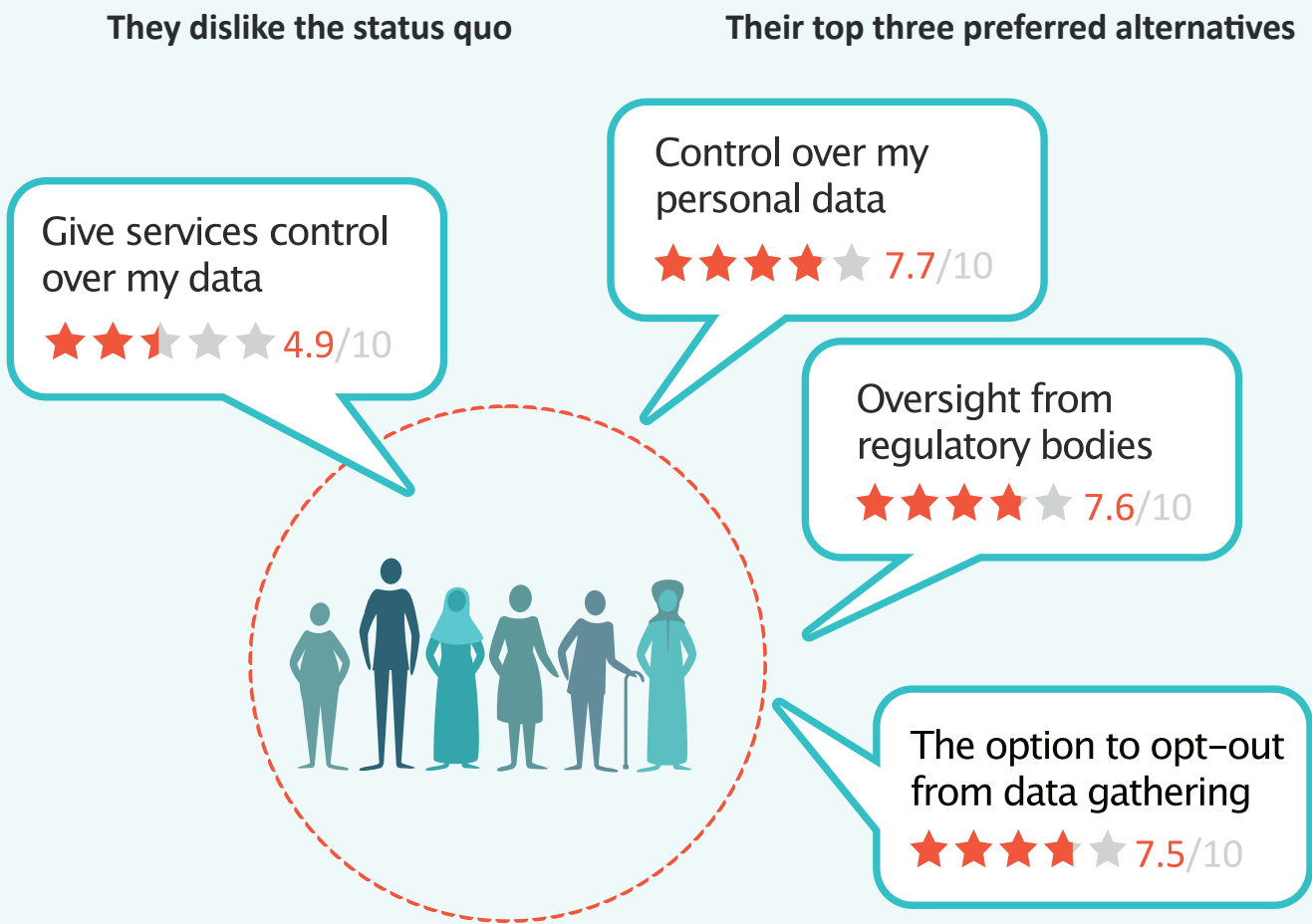
We conducted a nationally representative survey of 2,169 residents of the United Kingdom in May 2019 to understand their views on data management models.

We elicited views on eight models for managing personal data:

- Personal Data Store
- Responsible Independent Party
- Multiple Responsible Independent Organisations
- Digital Service (status quo)
- Data Co-operative
- Public Data Commons
- Regulatory Public Body
- Data ID Card (opt out option)

## KEY FINDINGS

### How the public rates models for managing their personal data



- A consistent finding is that respondents dislike the status quo, in which commercial organisations control personal data in return for the digital services they provide (average rating = 4.9/10, where 10 is excellent).
- Respondents generally preferred approaches that give individuals control over their personal data (average rating = 7.7/10), that include oversight from regulatory bodies (average rating = 7.6/10) or that enable opting out from data gathering (average rating = 7.5/10).
- Variations of data trusts were preferable to the status quo, but not as widely preferred as models involving personal control, regulatory oversight or the ability to opt out. The four ‘trust-like’ models that we presented were rated between 5.9/10 and 6.4/10.
- The public prefers all credible alternatives to the status quo.
- These findings were consistent across different methods used in the survey: asking respondents to rate models on a scale, choose a preferred model from a randomly generated pair, and choose a preferred scenario from a randomly generated pair.

## Differences amongst respondents

- Existing knowledge about issues relating to data was a significant predictor of preferences in relation to four models. More knowledgeable respondents preferred approaches that offered more control and/or oversight over personal data by a regulatory public body than less knowledgeable respondents who rated the status quo higher. While this effect was significant, it was relatively small. In other words, this mattered, but not a great deal.
- Age had a significant impact on evaluations of the status quo. Younger respondents rated the status quo higher than those who were older than 34.
- Apart from these two findings, there were no other clear differences in data management model evaluations by demographic subgroups within the sample.

## Recommendations

- Our findings suggest that new approaches to data management are urgently needed, because there is a strong desire from the public for an alternative to the status quo. These new approaches need to give individuals control over their personal data and include oversight from regulatory bodies.
- The implementation of such alternative data management models will require investment of resources, to support technical development, testing and iteration, and public consultation.
- In addition, to further advance understanding of public views about data management models, more research is needed. We need to understand why people prefer particular data management models and the extent of public understanding of differences across models.

# TABLE OF CONTENTS

<b>1. Introduction to the study</b>	<b>4</b>
<b>2. Information about the survey</b>	<b>5</b>
<b>3. Information about respondents</b>	<b>7</b>
<b>4. Results: public views on data practices</b>	<b>8</b>
<b>5. Results: public views on data management models</b>	<b>12</b>
<b>6. Conclusions and recommendations</b>	<b>21</b>
<b>7. About the research and report authors</b>	<b>22</b>
<b>8. References</b>	<b>22</b>
<b>9. Appendices</b>	<b>23</b>
Appendix 1.	23
Table: List of data management models presented to respondents (and used in the single-attribute conjoint experiment)	23
Appendix 2.	24
Table: Stimulus materials used in the multiple-attribute conjoint experiment	24

## 1. INTRODUCTION TO THE STUDY

This study investigates public attitudes towards different models for managing personal data. A number of recent studies have looked at public attitudes towards personal data, focusing on issues such as privacy and trust, but there has been little exploration of what the public think about alternative approaches to data management. This study fills that gap.

The UK government has identified both public trust in data and data mobility as critical enabling factors to the success of the UK's National Data Strategy (DCMS 2018). Under the General Data Protection Regulation (GDPR), individuals have new legal rights with regard to their personal data, including, for example, rights to access and portability. GDPR provides a legal imperative for developments in this area, and there is mounting evidence of growing experimentation in, for example, personal data management technologies such as Personal Data Stores (e.g. digi-me, Databox, Solid), data portability (Cntl shift 2018) and responsible data stewardship (ODI 2019).

These developments are taking place against a backdrop of growing concern about data collection and use. Research shows low levels of public trust when it comes to data practices (Doteveryone 2018), described elsewhere as a 'data trust deficit' (RSS 2014). Growing awareness of the data trust deficit, combined with high profile failures to protect people's personal data from exploitation or mis-use, has led to cross-societal consensus on the need for responsible data practices as articulated by, for example, the think tank Doteveryone, independent research and deliberative body the Ada Lovelace Institute, and the UK government's new Centre for Data Ethics and Innovation.

In this context, it is vital that public views are factored into new developments and expert debates and decisions that will shape the future of the data economy. This survey provides insights into public attitudes to data management models and associated data practices. It advances current understandings of what citizens perceive to be desirable when it comes to the management of their personal data.

## 2. INFORMATION ABOUT THE SURVEY

In May 2019, a total of 2,169 respondents living within the UK completed our online survey from an opt-in Qualtrics panel. In the survey, we examined what participants thought about eight models for managing personal data. Each model was based upon approaches to data management that were

being considered in various forms at the time of administering the survey, including Personal Data Stores, data trusts and data co-operatives. The models are summarised in Table 1 below and described in full in Table A1 in the Appendix.

**Table 1: List of data management models evaluated in the survey**

Personal Data Store	to give you personal control over your data, which you can manage in a secure way.
Responsible Independent Party	you have a say over what happens to your data, but are not personally responsible for looking after it.
Responsible Independent Organisation	have legal responsibilities to manage access to your data in ways that represent the interests of <i>all parties involved</i> .
Digital Service (Status Quo)	gives services control over your data (this is what usually happens now).
Data Co-operative	your data is managed collectively, by the people whose data is in the co-operative.
Public Data Commons	to make data accessible so everyone can benefit from it.
Regulatory Public Body	to ensure that personal data are collected, stored and used in legal and fair ways.
Data ID Card (Opt Out Option)	to give people the option of opting out of having their data collected.

Before rating the models, respondents first completed knowledge questions to gauge their familiarity with and understanding of concepts relevant to the survey—specifically, personal data, open data and the GDPR. Following these questions, we provided respondents with the answers to them, which also served to ensure that everyone began subsequent sections with the same general information about the topic. We also included questions that asked respondents to indicate their level of agreement with a number of statements, which were intended to assess their attitudes towards how personal data is collected, stored, used and shared by organisations.

Subsequently, we examined respondents' views toward each of the eight data management models using several different methods. Firstly, respondents were asked to rate four randomly selected data management models (presented one at a time) using a Likert scale ranging from 0 ('poor') to 10 ('excellent'). Secondly, because assigning a numeric value on an 11-point scale may have been difficult for some

respondents, we assessed their preferences through an innovative approach called a conjoint experiment (e.g. Hainmueller et al. 2014).

A conjoint experiment works by presenting respondents with profiles randomly generated from a list of selected characteristics. The decision task involves comparing the profiles side-by-side and then choosing the preferred option. A conjoint experiment is useful in this context because its forced choice design simplifies the decision task facing respondents (Hainmueller and Hopkins 2015, Pelzer 2019). We used a single-attribute conjoint experiment in which participants were presented with two randomly selected models from the list of eight and asked them to select the model that they preferred from the pair. This paired selection task was repeated three times. Our approach allowed us to evaluate how respondents rated the models independently and in comparison to one another. Table 2 provides an example of the single-attribute conjoint experiment used in this study.

Table 2. Example of the single-attribute conjoint experiment

Option A	Option B
You are given a secure place to collect, store and manage the data about you which has been collected by other services. This is called a Personal Data Store, or PDS. You have access to this data, and you can decide who else can access this data, how they can use it and under what circumstances. The purpose of the PDS is to give you personal control over your data, which you can manage in a secure way.	You are given a way to nominate a responsible independent party to oversee collection, storage and access of your personal data. They have legal responsibilities to look after your data. In line with your wishes, the nominated party can make decisions on your behalf about who accesses your data, what they can do with it and under what circumstances. You have a say over what happens to your data, but you are not personally responsible for looking after it.
Based on these descriptions, which option for managing data would you prefer? <input type="checkbox"/> Option A <input type="checkbox"/> Option B	

Our third method for assessing respondents’ views in relation to data management models was to ask them to complete a multiple-attribute conjoint experiment, in order to explore the importance of different characteristics that may influence their preferences with regard to data use, handling and management. We did this by randomly combining multiple features in profiles, to assess the relative effect of each characteristic on preferences for data management approaches. We asked respondents to express preferences for profiles of data scenarios generated from a combination of these factors:

- Type of data (for example, medical, financial, media consumption);

- Who has control (for example, individual, trustee, commercial organisation);
- What management arrangements mean for the individual (for example, full control over what happens to data, know what data is held about them, by whom and what they do with it);
- Use and beneficiaries of the data (for example, personal insights, generate profit, benefit society).

This multiple-attribute conjoint experiment enabled us to simultaneously explore the significance of different factors, which would be impractical in a typical survey experiment. The full details of this conjoint experiment are presented in Table A2 in the Appendix, and an example is provided in Table 3 below.

Table 3. Example of the multiple-attribute conjoint experiment

	Option A	Option B
In this scenario the data is	Medical data	Financial data
The data is controlled by	You	A trustee like a city council or the government
You will be able to	Have full control over what happens to it	Know what data is held about you, by whom and what they do with it
The data will be used for these reasons, and generate these benefits	So you can get insights and value from your personal data	So an organisation can use your data to benefit the public

Following the conjoint experiments, we asked several other questions about attitudes to personal data collection and use, data-related knowledge and online behaviours, and standard demographic

questions (about age, gender, ethnicity, etc), to further differentiate our analyses. The full survey can be accessed here: <https://livingwithdata.org/previous-research/views-on-data-management/>.

### 3. INFORMATION ABOUT RESPONDENTS

We collected data from a diverse group of respondents from across all regions of the UK, including Northern Ireland (for a full demographic breakdown, see Table 4). Quotas matched against UK Census demographics were used for gender and age, and there is also a good mix of educational attainment, employment status, and household income, enabling us to differentiate our analyses by various subgroups. Respondents were recruited by Qualtrics using opt-in methods, the sample demographics of which compare favourably with other reputable Internet panels such as the British Election Study conducted by YouGov (see column 2 in Table 4). As the survey was

conducted online using an Internet panel, it is likely that respondents were capable technology users. This was confirmed in answers to related questions. For example, 94.6% of respondents indicated that they were confident using devices to do things online, and 98.9% stated they used the Internet daily. The most common way of accessing the Internet was via mobile phone (83.9%). Only 8.5% of respondents indicated that they were not users of at least one of the major social media platforms (e.g. Facebook, Twitter, Instagram).

**Table 4. Respondent demographics compared to BES (%)**

		<b>This Sample: Qualtrics May 2019</b>	<b>Comparison: British Election Study March 2019</b>
<b>Gender</b>	Male	47.40	45.92
	Female	52.19	54.08
	Other (non-binary)	0.41	--
<b>Age</b>	18-34	32.64	17.00
	35-54	38.13	33.68
	55 or older	29.23	49.58
<b>Education</b>	No formal qualification	5.17	6.44
	Technical or other qualification	18.74	22.42
	GCSE/A-Level (or equivalent)	48.92	40.45
	University degree (or higher)	27.18	30.41
<b>Employment Status</b>	Full time	44.20	39.40
	Part time	16.91	15.25
	Not working	24.34	16.17
	Retired	14.55	5.85
<b>Household income</b>	Less than £15,000	21.20	14.06
	£15,000 to < £30,000	32.88	31.52
	£30,000 to < £50,000	26.16	27.74
	More than £50,000	19.76	22.03
<b>Ethnicity</b>	White	90.63	95.74
	BAME	9.37	4.26
<b>Disability</b>	Disabled	20.94	31.35
	Non-disabled	79.06	68.65
<b>Total %</b>		100.00	100.00
<b>N</b>		2,169	30,842

*Note: Our data was collected from members of a self-selected Internet panel by Qualtrics in May 2019. British Election Study (BES) data was collected by YouGov in March 2019. Respondents who provided a 'don't know' answer or refused to answer a question are not included in these totals. Not all percentages sum to 100 due to rounding.*

## 4. RESULTS: PUBLIC VIEWS ON DATA PRACTICES

### 4.1. Public knowledge of key concepts

We presented participants with a series of statements about personal data, open data and the GDPR and asked them to identify whether each statement was true or false. These statements were used to assess their knowledge about relevant issues and evaluate responses to later questions in light of this data. Some of these statements were reverse worded to account for potential agreement bias.

Respondents appeared most knowledgeable about the concept of personal data, with the vast majority correctly answering questions related to its definition (more than seven out of every 10 respondents got

these questions correct; see Table 5). Respondents were least knowledgeable about open data – less than half (ca. 48%) were able to correctly answer two questions on this topic. Results were mixed concerning familiarity with and understanding of the GDPR: 53% of the sample provided correct answers to a question about data portability and 93% correctly answered a question about the main purpose of the GDPR. Overall, the mean number of items correct in the full sample was 5.5 out of 8, which demonstrates a fair understanding of the concepts.

**Table 5. Percentage of knowledge questions answered correctly**

Question (correct response)	% Correct
The General Data Protection Regulation (GDPR) governs the processing of personal data (collection, storage, and use). <i>(True)</i>	93.1
Any information that can be used to identify an individual is personal data. <i>(True)</i>	92.2
Location data collected by your mobile phone is not personal data. <i>(False)</i>	73.4
The General Data Protection Regulation (GDPR) does not give you the right to access the personal data organisations hold about you. <i>(False)</i>	72.2
There are still no financial penalties for companies that do not comply with the General Data Protection Regulation (GDPR). <i>(False)</i>	69.0
The General Data Protection Regulation (GDPR) allows for 'data portability' meaning that you can take your data from one organisation and give it to another. <i>(True)</i>	52.6
Open data does not generally include personal data. <i>(True)</i>	48.9
Open data can only be used, modified, and shared for non-commercial purposes. <i>(False)</i>	48.2
<b>Mean number of correct items</b>	<b>5.5</b>



## 4.2. Public attitudes towards data practices

We asked respondents their views on data uses and potential future data-driven services. To assess views on data uses, we asked participants to indicate on a 5-point Likert scale whether they agreed or disagreed with a series of statements. Statements and responses can be seen in Table 6 below.

Clearly, respondents were concerned about the privacy (84.6% agreement) and security (84.2%) of their personal data, they wanted to be able to exercise their rights (92.1%) and have more control over their data (89.0%). In particular, they were concerned about how their personal data is used by organisations (86.9%), and they wanted companies to be held accountable if their personal data is being misused (96.1%).

Respondents were against commercial organisations using personal data to generate profit (78.3%). Only around half of respondents supported sharing personal data for use in research in the public interest (52.7%). Around two in three want data to be used for the social good (68.8%). Most want data to be managed, analysed and gathered in ethical ways (84.0%).



**Table 6. Percent agreement with statements about data collection and use**

Statement	% Agree
I want companies to be held accountable if they misuse my personal data.	96.1
I want clear, easy-to-find explanations about what happens to my data when I share it online.	92.2
I want people to be able to exercise their rights over their personal data.	92.1
I want more control over my personal data.	89.0
I want more control over how my personal data is used by organisations	86.9
I'm concerned about social media companies selling my personal data to advertisers.	85.6
It should be difficult for organisations to collect sensitive personal data (e.g. genetic information).	84.8
I'm concerned about the privacy of my personal data online.	84.6
I'm concerned about the security of my personal data online.	84.2
I want data to be gathered, analysed and managed in ethical ways.	84.0
I want collective approaches to data that work for everyone so that the most vulnerable people in society are protected.	81.2
I'm against corporate profit-making from personal data.	78.3
I'm concerned about services using my personal data to generate personalised recommendations.	74.3
I believe that businesses should publish open data in the same way as governments.	69.4
I would rather companies only use open data instead of personal data.	69.4
I want data to be used for the social good.	68.8
I believe in making data available for science, research or technical experimentation in ways that benefit society.	67.3
I want an alternative model for managing personal data to what currently exists.	62.0
I want to profit personally from my own data.	60.7
I don't understand what happens to my personal data when I share it online.	60.0
I'm comfortable with any organisation using information about me so long as it has been anonymised first.	59.2
I support sharing my personal data when it's used for research in the public's interest.	52.7
I believe in the social benefits of gathering, aggregating and analyzing data.	51.9
I support sharing my personal data with my energy company if it means getting more accurate assessments of my energy use.	51.1
I'm in favour of personal data being collectively managed by a community of users who the data is about.	40.2
I'm in favour of open data.	39.3
I don't mind sharing my personal data with commercial organisations to help them develop new products.	37.5
I believe in making data available for science, research or technical experimentation for commercial purposes (i.e. making a profit).	34.8
I don't have strong opinions about personal data collection and use.	21.5

To gauge respondents' views about the types of data-driven apps and services that they would like to see developed in the future, we asked respondents to select services from a list or add their own. Listed services and responses are shown in Table 7 below.

The most popular were in health and wellbeing (66.3% of the sample selected this type of service), the environment (53.6%), and schools and education (50.8%). The least popular were entertainment services and quizzes (16.2%).

**Table 7. Percent of sample selecting each data-driven service type**

Type of service	% Selected
<b>Health and wellbeing</b> (for example, about physical and mental health and support)	66.3
<b>Environment</b> (for example to help individuals or communities reduce their carbon footprints)	53.6
<b>School and education</b> (for example, supporting children, teachers and parents)	50.8
<b>Finance and budgeting</b> (for example, to help with everyday money matters)	44.9
<b>Education and career</b> (for example, to support you in further education and career paths)	43.3
<b>Local community</b> (for example, bringing people together around local events and issues)	43.2
<b>Family services</b> (for example, aimed at helping with day to day family life)	41.5
<b>Learning and skill development</b> (for example, learning a new language)	39.5
<b>Everyday life</b> (for example, advice and support for real life scenarios)	36.2
<b>Citizen science</b> (for example, taking part in activities that help solve big issues)	33.9
<b>Politics</b> (for example, telling you how your MP voted or how new policies affect your area)	33.2
<b>Sport/fitness and healthy eating</b> (for example, to help with fitness and eating well)	30.9
<b>Travel</b> (for example, learning about places and culture and tips for experiencing the world)	29.5
<b>Communication</b> (for example, to connect you with friends and family)	27.6
<b>Personal goals</b> (for example, setting targets, getting support and tracking progress)	23.1
<b>Accessing TV and music</b> (for example, services like iPlayer, Netflix and Spotify)	22.8
<b>Home-based DIY</b> (for example, cooking, gardening, crafts)	22.6
<b>Archival</b> (for example, making or exploring archives of images or videos)	22.0
<b>Debate/discussion</b> (for example, to explore your views and the views of others)	19.0
<b>Entertainment and quizzes</b> (for example, entertaining games or personal quizzes)	16.2

We also asked respondents who they would like to see provide these services. Most respondents said they would like to see governmental or publically-funded organisations provide such services – 46% and 40% of respondents selected these options, compared to 18% selecting commercial organisations in a question where respondents could select as many options as they wanted.

# 5. RESULTS: PUBLIC VIEWS ON DATA MANAGEMENT MODELS

## 5.1. Rating different management models

Most companies collect data about their users' online behaviours. These companies have significant control over the data, where it is stored and how it is used. However, other approaches to data collection and storage are being considered, which some people believe would be better for individuals and society. We examined respondents' attitudes towards these alternative approaches to data management, including various iterations of data trusts, Personal Data Stores, the status quo, and an 'opt-out' option where respondents could indicate that they did not want their data to be collected.

Of the eight data management models that we presented to respondents, the most preferred approach was the Personal Data Store (or PDS), described in the survey as 'a secure place to collect, store and manage the data about you which has

been collected by other services' (see Table 8 for mean ratings of each model). The Personal Data Store would give individuals control over their personal data. Responses to questions about views on data uses, reported in Table 6 above, suggest that this may be why this model was highly rated: 86.9% of respondents agreed with the statement 'I want more control over how my personal data is used by organisations', and 89.0% agreed with the statement 'I want more control over my personal data'. Previous research has also emphasized the importance of control – e.g. 94% of participants in a 2015 Digital Catapult survey said they wanted more control over their data (Digital Catapult 2015).

**Table 8. Mean ratings on a scale from 0 to 10 for each data management model**

Model	Mean Rating
Personal Data Store	7.7
Regulatory Public Body	7.6
Data ID Card (with clear opt-out options)	7.5
Responsible Independent Organisations	6.4
Public Data Commons	6.3
Responsible Independent Party	6.2
Data Co-operative	5.9
Digital Service (Status Quo)	4.9

After the PDS, the next highest rated model was one involving oversight by a regulatory public body which, the survey explained, would oversee ‘how organisations access and use data, acting on behalf of UK citizens’ in order to ‘ensure that personal data are collected, stored and used in legal and fair ways’. As noted above, elsewhere in the survey, we asked respondents who they would like to see provide new data-driven services ‘for the public good’ and most selected governmental or publically-funded organisations (46% and 40% of respondents respectively). This reinforces the finding that oversight of data-related developments by a public regulatory body was a strong preference for our respondents.

The high rating of this model by respondents suggests a preference for legally enforceable safeguards alongside the personal control of data offered by the PDS. This finding was reinforced in responses to questions about views on data uses, discussed above, in which:

96.1% of respondents agreed with the statement: *‘I want companies to be held accountable if they misuse my personal data’*.

92.2% of respondents agreed with the statement: *‘I want clear, easy-to-find explanations about what happens to my personal data when I share it online’*.

84.8% of respondents agreed with the statement: *‘I think it should be difficult for organisations to collect sensitive personal data (e.g. genetic information)’*.

Realising these statements – for example, holding companies accountable – requires governance and this explains respondents’ strong preference for a data management model overseen by a regulatory public body. A Royal Statistical Society survey in 2014 found that “on balance, the research suggests there is more support for the government preventing misuse of personal data than there is an appetite to have personal control over this” (2014, p.3). In contrast, we found a strong preference for governance and personal control. The high rating of both the PDS and oversight by a regulatory public body suggests that both personal control and oversight are important principles in data management for respondents. Both approaches would result in uses of data that are preferable to the status quo.

One of the models that we presented to respondents would allow people to opt out of having their data collected. We described this as a ‘Data ID Card’, to give material form to a model for opting out of data collection. The relatively high rating of this model (third overall, out of all 8 models) reinforces the importance of individual control over data amongst our respondents, as well as indicating a strong dislike of the status quo.

Data trusts, defined by the ODI (2018) as ‘a legal structure that provides independent stewardship of data’ for the benefit of all parties. Data trusts, co-operatives and commons-based models have a number of differences, but all involve trusted parties overseeing, managing and stewarding data on behalf of individuals and communities. In this sense (rather than in a legal sense), they are all trust-like. In our survey, we explored all three of these models:

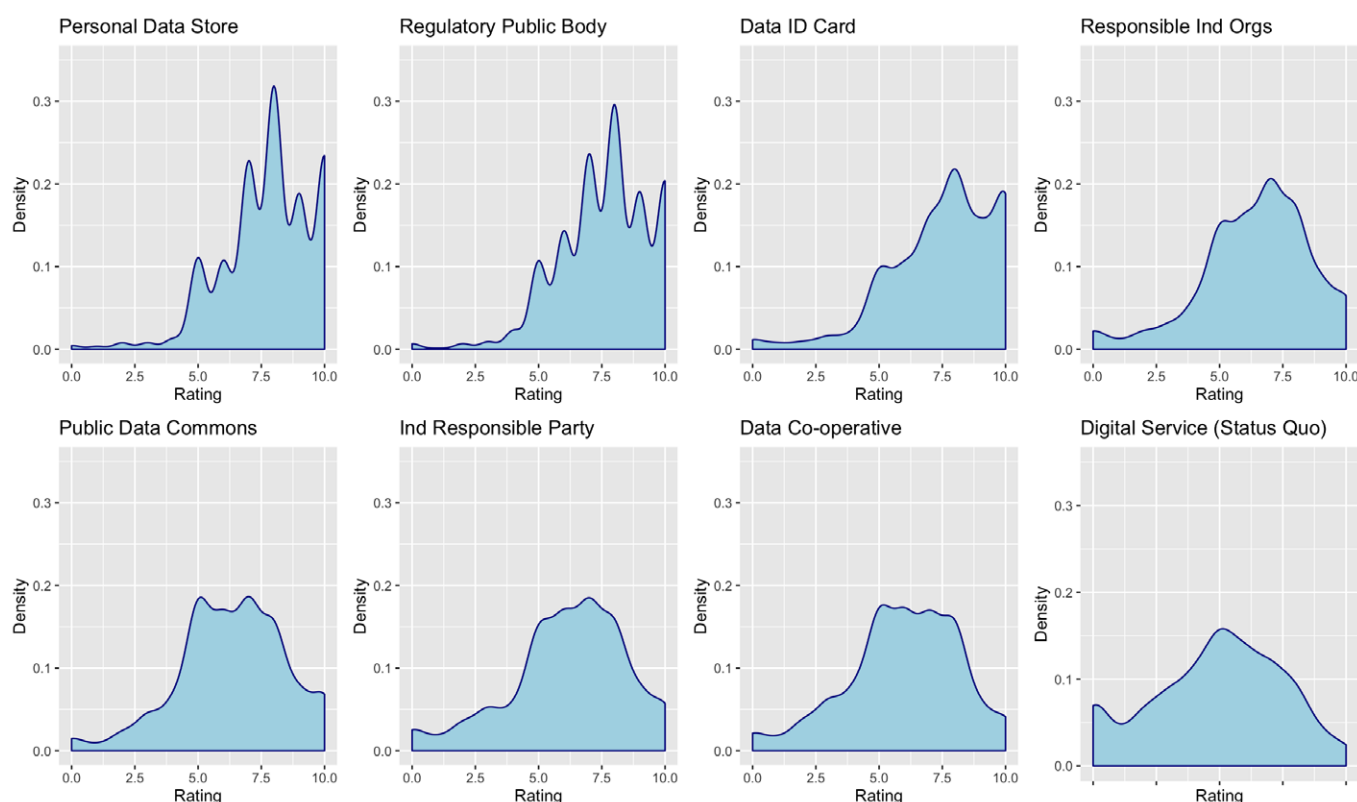
- the data co-operative, which manages the collection and storage of its members’ data, is accountable to its members and is governed by a board of representatives constituted by its members;
- the data commons, similarly collectively motivated, which enables online access to community data which can be used for various purposes and for the benefit of all; and data trusts;
- two types of trust:
  - a trust governed by a responsible independent party which makes decisions on behalf of data subjects about who accesses data, what they can do with it and under what circumstances,
  - a trust governed by multiple responsible independent organisations which manage different types of data in different contexts (for example, one for health data, one for finance data, and so on) and represent the interests of all parties involved.

These four trust-like models were preferable to the status quo, but not as widely preferred as models involving personal control, regulatory oversight or the ability to opt out. They all had lower mean scores than those that offer personal control or regulatory oversight. These models may have received lower ratings because they were less familiar to respondents than models based on the more commonplace concepts of personal control and regulation. As noted above, in the knowledge questions with which we opened the survey, respondents demonstrated limited knowledge of open data. Elsewhere in the survey, only 39.3% of respondents agreed with the statement ‘I’m in favour of open data.’ This relatively low level of support for open data could result from the equally low levels of understanding that we identified, and

together, these findings may explain the lower mean scores for the ‘trust-like’ data management models that we presented to respondents, which build on some of the principles of open data.

Figure 1 provides additional information regarding the ratings of the different models. The figure displays the distribution of ratings for each approach. It shows that high mean ratings for the top three approaches – PDS, oversight by a regulatory public body, and opting out – occur because ratings for these models are overwhelmingly positive, with very few ratings under the midpoint on the scale. In contrast, responses to the other approaches to managing data have more ‘bell-curve’ shaped distributions, reflecting a broader range of opinions about them, some positive, and some negative.

**Figure 1. Distributions of ratings across data management models**



The models that did not offer personal control or governmental/regulatory oversight had lower mean scores than those that did, including models overseen by a public data commons (which ‘collects, stores and manages access to open data which can be used for various purposes’), a data co-operative (which ‘manages the collection and storage of its members’ data and is accountable to its members’), responsible independent organisations (which ‘manage your data in different contexts’ and ‘make decisions about who can access your data, what they can do with it and under what circumstances’) or a responsible independent party (which ‘oversees collection, storage and access of your personal data’). It is possible that these models received lower ratings because they were less familiar to respondents. As noted above, in the knowledge questions with which we opened the

survey, respondents demonstrated limited knowledge of open data. Elsewhere in the survey, only 39.3% of respondents agreed with the statement ‘I’m in favour of open data’. Together, these findings seem to suggest relatively low levels of understanding of commons-based approaches to managing data, and this might explain these lower mean scores.

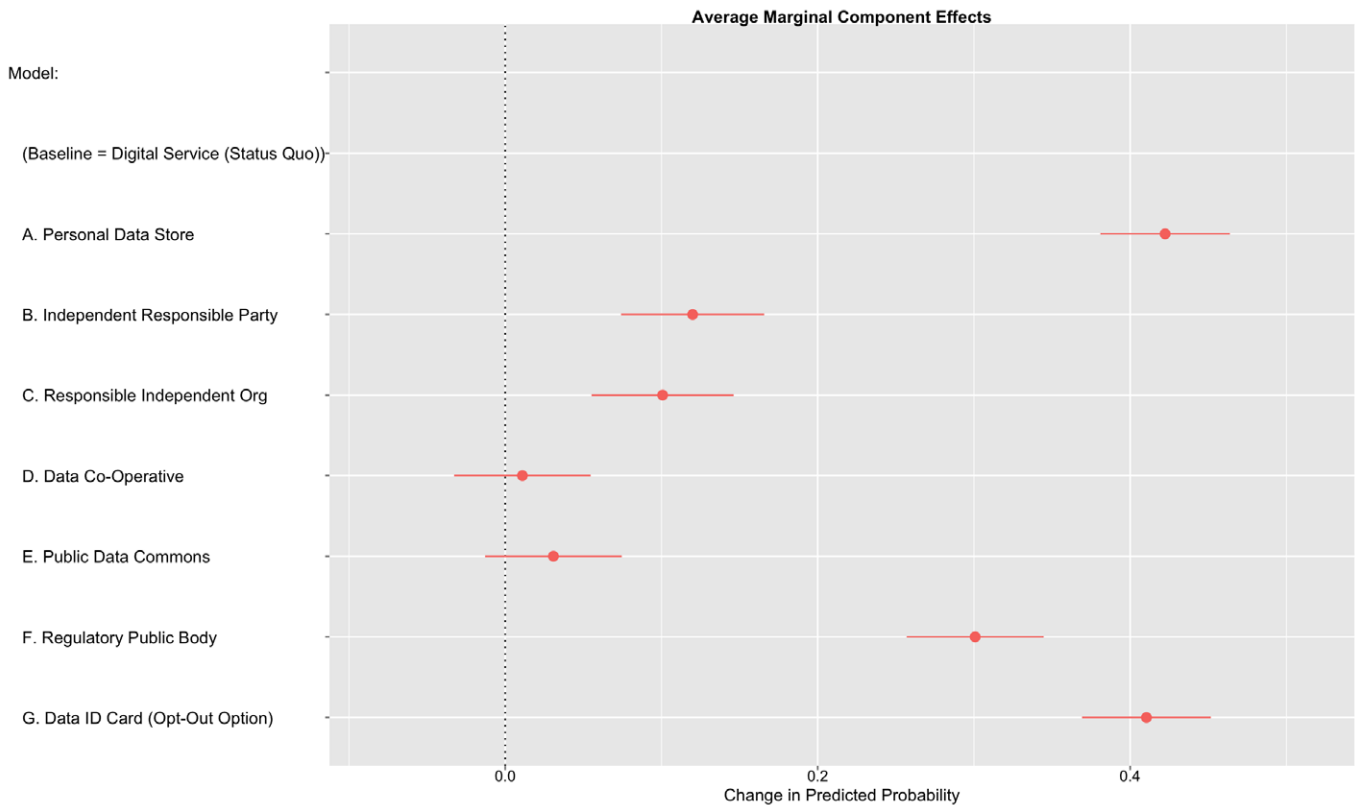
What is most striking about the results is that respondents preferred *all other models to the status quo*, described in the survey as a ‘digital services model’ that ‘gives services control over what happens to your data’, the prevailing online model at the time of writing. With an average rating of 4.9 out of 10, this suggests that respondents are not happy with the current ways in which services and organisations control data.

## 5.2. Comparing different data management models

The results of the single-attribute conjoint experiment also corroborate these findings. The conjoint experiment presented respondents with a randomly generated pair of data management models, and respondents were asked to choose the option that they preferred. The conjoint analysis is presented in Figure 2. The plotted points provide the change in the probability of selecting a model relative to the status quo (i.e. digital services having control over people’s data). These plotted points are known as average marginal component effects (AMCEs). The vertical dotted line indicates this baseline; points to the right of the dotted line indicate an increase in the probability of choosing that particular model relative to the baseline. The lines around each side of plotted points are error bars, indicating uncertainty around each value, which derives from the fact that our survey is based on a sample of the population, not the whole population.

As with the individual ratings task, the conjoint analysis reveals that the top three preferred models are the PDS, opting out and oversight by a regulatory public body, in that order of preference. There was at least a 30 percentage point increase in selecting the top three data management models compared to the current, status quo approach. In other words, respondents were at least 30 percentage points more likely to choose the top three models. This is a significant number, both statistically and substantively. Also as with the rating task, trust-like models were preferable to the status quo, but less preferable than those based on personal choice and control and regulation.

Figure 2. Results from the single-attribute conjoint analysis



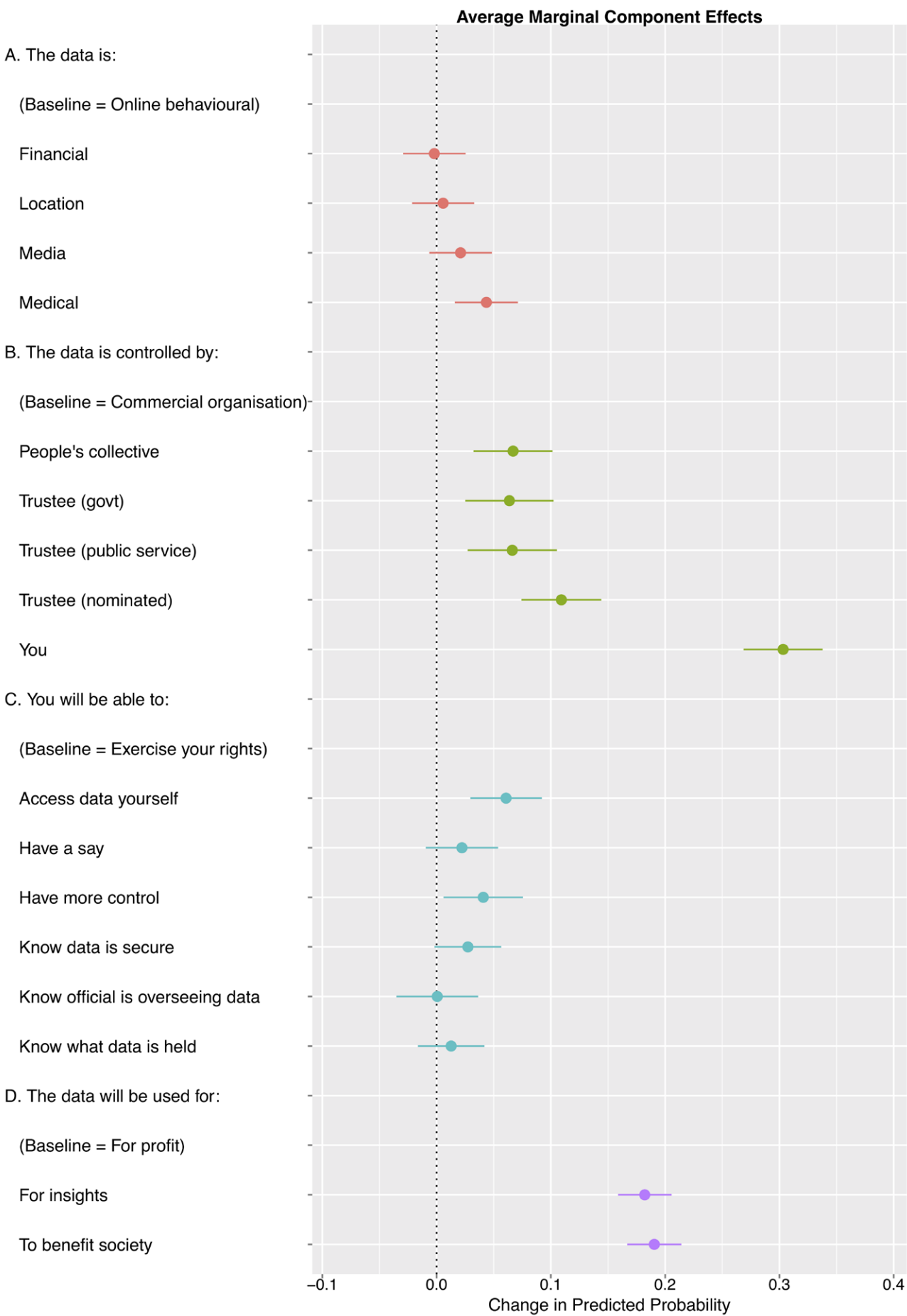
5.3. Preferences in relation to data handling scenarios

As noted in the introduction, we asked respondents to complete a multiple-attribute conjoint experiment, in order to explore the importance of different characteristics that may influence their preferences with regard to data use, handling and management. We discuss the results of this experiment in this section.

Figure 3 displays the results from the multiple-attribute conjoint analysis, which compared the significance of a number of attributes in data handling scenarios, such as types of data, control and rights, uses of data and related benefits. As with the single-attribute conjoint analysis, in the figure we present results that show the change in the probability of selecting a profile with particular characteristics relative to a baseline, this time for each of the attributes we included in the scenarios.



Figure 3. Results from the multiple-attribute conjoint analysis



The single most important factor influencing responses is the locus of control over the data. When presented with data management scenarios that give respondents control over their own data, then the probability of selecting that profile increases by 0.30 relative to the baseline (i.e. a commercial organisation controls the data). Thus when the data is controlled by 'you' instead of a commercial organisation, we see the largest increase in the probability of selecting that profile. In other words, controlling their own data was really important to respondents.

As noted above, control seems to be an important consideration in evaluating different approaches to managing personal data (as seen in Table 8), as well as in responses to various statements (as seen in Table 6). It is therefore unsurprising that it plays a key role in this analysis. As we discovered in responses to other questions, respondents preferred scenarios in which anyone other than a commercial organisation was responsible for controlling the data, though there was little notable differentiation among the alternative controllers that we presented in this scenario, apart from respondents themselves, for which respondents expressed a significant preference.

Other significant attributes relate to uses and beneficiaries of the data. Respondents were more likely to prefer scenarios in which data would be used for personal insights or to benefit society than for profit. The effect sizes for these factors are in the medium range (a change in the probability of selecting that profile of 0.15 or greater). In other words, there is a 15-percentage point increase in the chance that a particular profile would be selected when it provided personal insights or benefits to society compared to for profit.

Finally, respondents preferred scenarios that provided them with the right to access their personal data, have more control over it, and know that it is secure. But compared to other factors discussed above, individual rights contributed only small effects to the overall decision, with the change in the predicted probabilities of less than 0.1.

## 5.4. Differences amongst respondents

We examined how opinions about the data management models might vary amongst different subgroups of respondents, according to demographic characteristics and knowledge of related matters.

Existing knowledge about issues relating to data was a significant predictor of preferences in relation to four models. More knowledgeable respondents preferred approaches that offered more control and/or oversight over personal data by a regulatory public body than less knowledgeable respondents who rated the status quo higher. This effect was relatively small (about a half point difference on a 10-point scale).

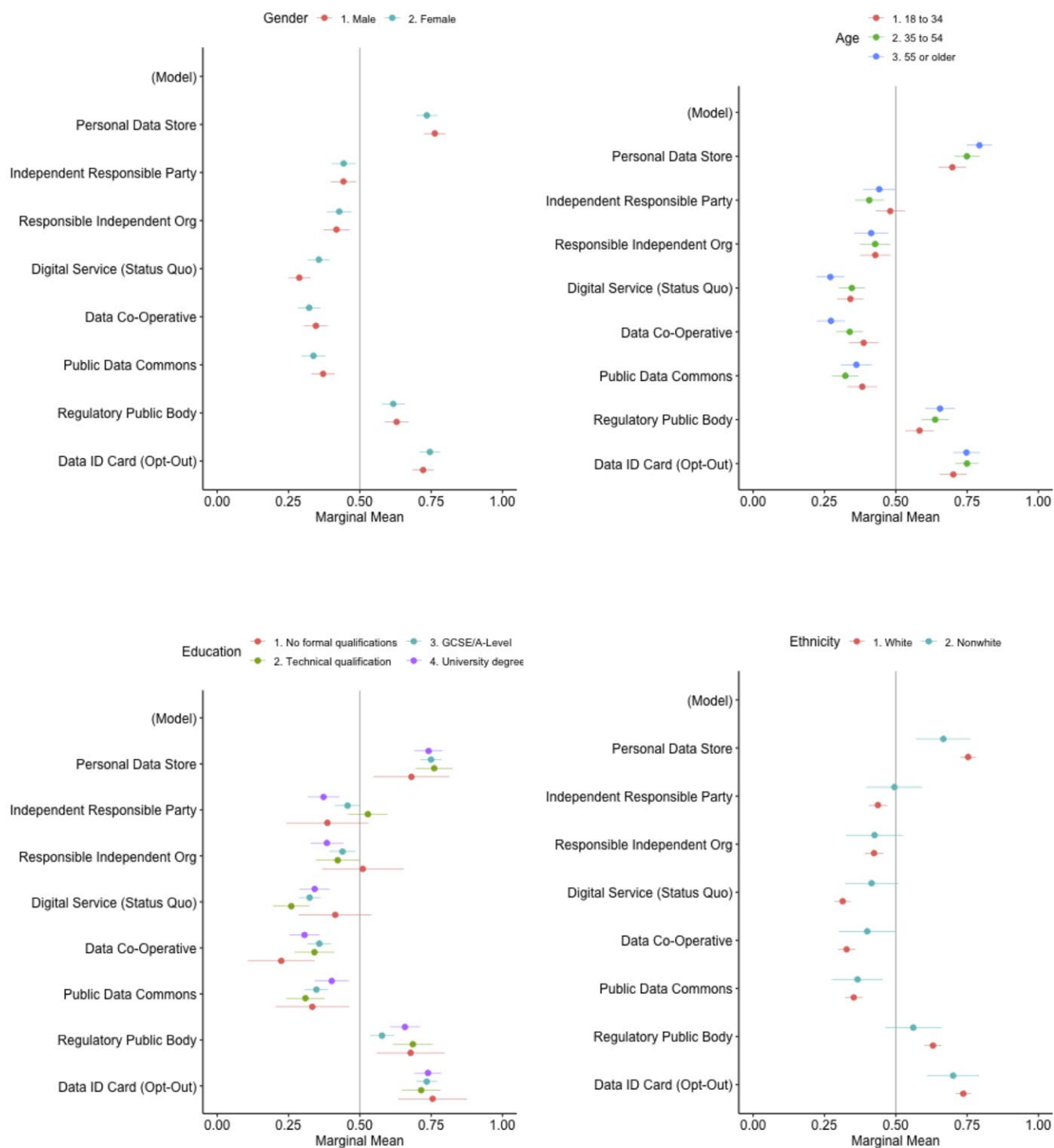
Age also had a significant impact on evaluations of the status quo approach: younger respondents rated the status quo model higher than those who were older than 34 (the effect size amounts to a about 1 point higher mean ratings on a 10-point scale). In other words, differences relating to age and existing knowledge mattered, but not a great deal.

Apart from these two findings, there were no other clear differences in data management model evaluations by demographic subgroups within the sample.

In addition, we explored subgroup differences for the single-attribute conjoint experiment, as seen in Figure 4. This figure plots the average proportion of respondents selecting each data management model, also known as marginal means. By design, marginal means average 0.5. In other words, if responses were simply randomly chosen, there is a 50:50 chance that a given response is selected.

Values above 0.5 tell us that respondents prefer a given model, and values below 0.5 indicate that respondents do not like the model. A value of 0 would tell us that the model was never selected; a value of 1 means that it was always selected. As with previous figures, Figure 4 also includes error bars.

**Figure 4. Subgroup analysis of the single-attribute conjoint analysis**



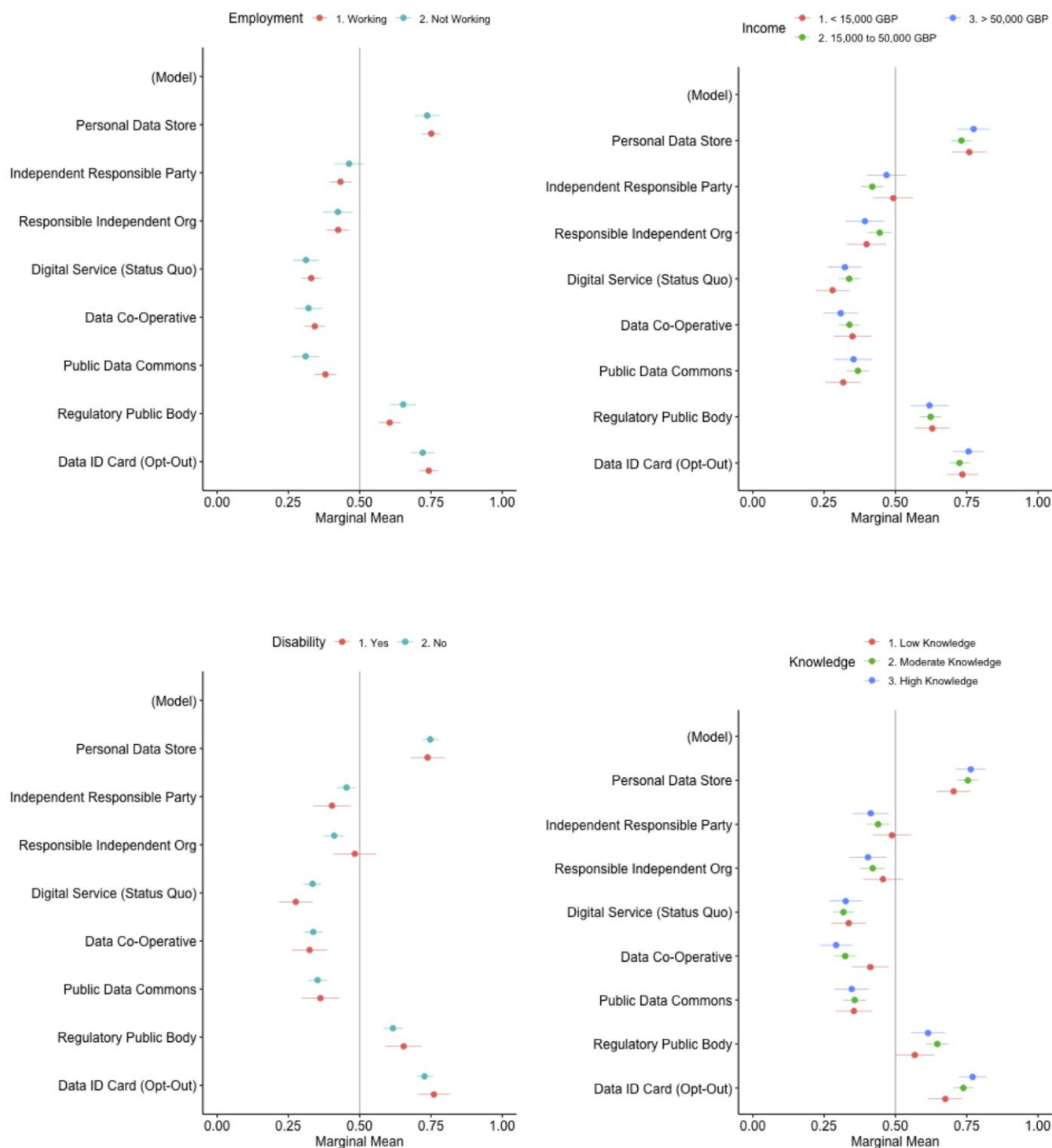


Figure 4 demonstrates that there are few large differences by subgroup or demographic characteristics in selecting a data management model – the plot points for the various subgroups are, for the most part, grouped closely together. However, there are some exceptions. For instance, age appears to have some influence on the models chosen by respondents. Young people appear less swayed by

many of the popular data management models (such as the PDS, oversight by a regulatory public body, and the opt out option). This is indicated in Figure 4 by the closer proximity to the 0.5 value for younger respondents. Less knowledgeable respondents, in general, were less likely to differentiate among any of the models. Their marginal means were also closer to the 0.5 vertical line.

## 6. CONCLUSIONS AND RECOMMENDATIONS

Our findings suggest that new approaches to data management are urgently needed, because there is a strong desire from the public for an alternative to the status quo. A consistent finding is that respondents dislike the status quo, in which commercial organisations control personal data in return for the digital services they provide. Respondents generally preferred approaches that give individuals control over their personal data, that include oversight from regulatory bodies or that enable opting out from data gathering. When a range of credible alternatives are available—for example, a public data commons, a data co-operative, oversight by a responsible independent organisation or party—respondents preferred all of these approaches to data management to the status quo. Existing knowledge and age had an impact on evaluations of models, but the effects of these variables was relatively small.

These findings were consistent across different methods used in the survey: asking respondents to rank models on a scale, choose a preferred model from a randomly generated pair, and choose a preferred scenario from a pair made up of randomly generated features. Policy-makers and organisations that handle personal data need to accept that the status quo is not sustainable—that is if they wish to factor public views into new developments and decisions that will shape the future of the data economy.. New approaches need to give individuals control over their personal data and include oversight from regulatory bodies.

However, not all alternative approaches to data management were rated equally. Respondents expressed a greater preference for some than for others. Data trust-like models—a public data commons, a data co-operative, oversight by a responsible independent party or organisations—were ranked below PDS, regulatory and opt out models.

These findings were consistent across different methods used in the survey. We cannot therefore conclude that there is a ‘huge appetite’ for data trusts amongst the public, as the ODI concluded from their data trust pilots with organisational stakeholders (2019). Further research is needed to explore the reasons for this finding.

The implementation of such alternative data management models will require investment of resources, to support technical development, testing and iteration, and public consultation. In addition, to further advance understanding of public views about data management models, more research is needed. We need to understand why people prefer particular data management models and the extent of public understanding of differences across models.

In short, for there to be public support for data practices and data-driven developments, it is clear that:

1. **New approaches to data management are urgently needed**, which give individuals control over their personal data and include oversight from regulatory bodies.
2. **Investment of resources is required**, to support technical development, testing and iteration, and public consultation, if the implementation of alternative data management models is to be successful.
3. **More research is needed** to further advance understanding of public views about data management models.

## 7. ABOUT THE RESEARCH AND THE REPORT AUTHORS

The research formed part of an ongoing collaboration between the University of Sheffield's Digital Society Network (<https://www.sheffield.ac.uk/faculty/social-sciences/digital-society-network/home>) and BBC R&D (<https://www.bbc.co.uk/rd>). It was made possible by funding from the Arts and Humanities Creative Economy Engagement Fellowship Scheme (<https://ahrc.ukri.org/innovation/knowledgeexchange/cultural-engagement-fund/>), BBC R & D and the University of Sheffield.

A shorter version of the research report can be found here: <https://livingwithdata.org/previous-research/views-on-data-management/>.

Helen Kennedy is professor of digital society at the University of Sheffield, director of the Digital Society Network and of the Living With Data programme of research hosted there (<https://livingwithdata.org/>).

**Todd Hartman** is senior lecturer in quantitative social science and a political psychologist, with international expertise in survey research and experimental design, also at the University of Sheffield.

**Robin Steedman** is a postdoctoral researcher interested in diversity and inequality in the media, formerly of the University of Sheffield and now of Copenhagen Business School.

**Rhianne Jones** is a research lead at BBC Research & Development. Her research focuses on examining the development and use of data-driven technology and the public interest.

## 8. REFERENCES

Department for Digital, Culture, Media & Sport (DCMS), 2018. Data mobility: The personal data portability growth opportunity for the UK economy. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/755219/Data\\_Mobility\\_report.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/755219/Data_Mobility_report.pdf) (accessed: April 2019).

Digital Catapult. (2015). Trust in personal data: A UK review. Report by Digital Catapult.

Hainmueller, J. and Hopkins, D.J., 2015. The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants. *American Journal of Political Science*, 59(3), pp.529-548.

Hainmueller, J., Hopkins, D.J. and Yamamoto, T., 2014. Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis*, 22(1), pp.1-30.

Knudsen, E. and Johannesson, M.P., 2019. Beyond the limits of survey experiments: How conjoint designs advance causal inference in political communication research. *Political Communication*, 36(2), pp.259-271.

Open Data Institute (ODI), November 20, 2018. UK's first 'data trust' pilots to be led by the ODI in partnership with central and local government. Available at: <https://theodi.org/article/uks-first-data-trust-pilots-to-be-led-by-the-odi-in-partnership-with-central-and-local-government/> (accessed: April 2019).

Open Data Institute (ODI), 15th April 2019. Huge appetite for data trusts. Available at: <https://theodi.org/article/huge-appetite-for-data-trusts-according-to-new-odi-research/>, 15th April 2019.

Pelzer, E., 2019. The Potential of Conjoint Analysis for Communication Research. *Communication Research Reports*, 36(2), pp.136-147.

Royal Statistical Society (2014) Trust in data and attitudes toward data use/ data sharing (2014). Available at: <https://www.statslife.org.uk/images/pdf/rss-data-trust-data-sharing-attitudes-research-note.pdf> (accessed 24 February 2019).



# 9. APPENDICES

## Appendix 1.

**Table: List of data management models presented to respondents (and used in the single-attribute conjoint experiment)**

Name	Description
Personal Data Store	You are given a secure place to collect, store and manage the data about you which has been collected by other services. This is called a <b>Personal Data Store, or PDS</b> . You have access to this data, and you can decide who else can access this data, how they can use it and under what circumstances. <b>The purpose of the PDS is to give you personal control over your data, which you can manage in a secure way.</b>
Responsible Independent Party	You are given a way to nominate a <b>responsible independent party</b> to oversee collection, storage and access of your personal data. They have legal responsibilities to look after your data. In line with your wishes, the nominated party can make decisions on your behalf about who accesses your data, what they can do with it and under what circumstances. <b>You have a say over what happens to your data, but you are not personally responsible for looking after it.</b>
Responsible independent organisations	<b>Responsible independent organisations</b> manage your data in different contexts (eg one for health data, one for finance data, etc). These organisations make decisions about who can access your data, what they can do with it and under what circumstances. <b>They have legal responsibilities to manage access to your data in ways that represent the interests of all parties involved.</b>
Digital Service (Status Quo)	You sign up to a new <b>digital service</b> (eg an online shop) that collects and uses your data. You are asked to agree to terms of use and a privacy policy beforehand. These describe how the service will collect, store and manage data about you. You are given settings you can alter, but you are not able to change or negotiate these terms or see how your data is used. <b>This approach gives services control over your data (this is what usually happens now).</b>
Data Co-operative	You become a member of a <b>data co-operative</b> that manages the collection and storage of its members' data and is accountable to its members. As a member, you can put yourself forward to sit on a board of representatives and make decisions about who has access to members' data, how it is used and under what circumstances. Or you can vote for other co-operative members to do these things. <b>The purpose of the data co-operative is that your data is managed collectively, by the people whose data is in the co-operative.</b>
Public Data Commons	You access data online about your area and community using an open data platform that is accessible to all citizens under commons law. This is called a <b>public data commons</b> . The data commons collects, stores and manages access to open data which can be used for various purposes. Everyone can access and use this data, in line with the commons' rules of engagement. <b>The purpose of the public data commons is to make data accessible so everyone can benefit from it.</b>
Regulatory Public Body	You have been given the details of a new regulatory public body that oversees how organisations access and use data, acting on behalf of UK citizens. This public body provides oversight over how organisations collect, store and use personal data. It can hold organisations accountable for misuse (eg fine organisations when they breach terms of use). <b>The purpose of the regulatory body is to ensure that personal data are collected, stored and used in legal and fair ways.</b>
Data ID Card (Opt out)	You have the ability to choose whether to opt out of online data collection, storage and use – this is called managing your data preferences. Your data preferences are stored on a <b>data ID card</b> . You can use this card to log onto online sites. The card automatically opts you out of data collection, storage and use according to your preferences and whenever this is possible. <b>The purpose of the data ID card is to give people the option of opting out of having their data collected.</b>

## Appendix 2.

**Table: Stimulus materials used in the multiple-attribute conjoint experiment**

In this scenario the data is	Online behavioural data (eg your social media interactions or Amazon purchases) Medical data (eg information about you from your GP) Media data (eg your viewing history on Netflix) Financial data (eg your pension and benefits details) Location data (eg your geographical movements, collected by your mobile phone)
The data is controlled by	You A trustee you choose to nominate A trustee like a city council or the government A trustee like a public service organisation A people's collective you are a member of A commercial organisation or service provider
You will be able to	Know what data is held about you, by whom and what they do with it Access the data yourself Have a say in who to share data with Exercise your rights (e.g. for portability, erasure or rectification) Have more control over what happens to it Know that it is secure Know that someone official is overseeing the data so it's not misused
The data will be used for these reasons & generate these benefits	So you can get insights and value from your personal data So profit can be generated by offering services from your personal data So your data can be used in ways that you feel benefits society



