Department for
Business, Energy
& Industrial Strategy

Companies House

VALUING THE USER BENEFITS OF COMPANIES HOUSE DATA

Report 3: Intermediaries

BEIS Research Paper Number 2019/015

September 2019

Any enquiries regarding this publication should be sent to us at:
enquiries@beis.gov.uk

Contents

1 Introduction

This report has been prepared by ICF Consulting Services Limited (ICF) in collaboration with Economics for the Environment Consultancy (eftec). It is part of a series of reports that presents the findings of research commissioned by the Department for Business, Energy and Industrial Strategy (BEIS) to value the user benefits of Companies House (CH) data.

1.1 Outline

This report (Report 3) estimates the value of CH data to private sector businesses that use CH data as an input to their own commercial products and services. This sub-group of users has been classified as 'intermediaries' as they are adding value to CH data and selling it on to end-users of the data.

The other reports in the series are described below:

- Report 1 sets out the methodological framework for the study.
- Report 2 presents willingness to pay (WTP) estimates for all users of CH data and provides a policy example of how this analysis can be used.
- Report 4 presents findings for another subset of users: ‘providers of public goods’ (defined as the public sector and other organisations that use CH data to deliver public goods and public benefits).
- A separate policy summary summarises the results of the research and draws overall conclusions about the value of CH data to users.

1.2 Total population of intermediaries

Report 2 explains that the profile of the overall population of users of CH data is not known due to incomplete information. This also means that the size of the population of intermediaries is also unknown.

Some evidence is available from the WTP survey that was undertaken as part of this study. The WTP survey received 85 responses from private sector businesses who stated that one of the main benefits of CH data for their organisation was *“information/data that we include in the products and services we sell to our customers”*. These 85 intermediaries represented 14% of the total number of survey responses, although it is not possible to use these findings to estimate the wider population of intermediaries due to the potential sample bias in the WTP survey.

Further evidence is available for users of CH's 'bulk' data products. These products provide large quantities of CH data that private sector users can use as an input into their own products and services. Users are required to register in order to gain access to the bulk data and tend to have relatively close working relationships with CH. CH provided anonymised information on all of these users to inform this study, which showed that there were 132 private sector businesses accessing bulk data products from CH at the time of the research.

Qualitative research was also undertaken with a sample of 15 of the 132 intermediaries who were accessing bulk data products. There was no duplication between the 15 businesses included in the qualitative research and the 85 intermediaries who responded to the WTP survey. However, it was not possible to determine whether there was any overlap between the 85 intermediaries who responded to the WTP survey and the other intermediary users of bulk data that did not participate in the research. It is therefore only possible to say that the total population of intermediaries includes at least the 132 private sector businesses that were accessing bulk data at the time of the research.

1.3 Methodology

The methodological framework for this study is described in more detail in a separate, supplementary report (Report 1). The research with 'intermediaries' involved fieldwork with the two groups described above:

- Intermediaries who responded to the quantitative survey of all users of CH search and data services. The survey explored users' WTP for CH data and the overall results are presented in Report 2, based on all 608 responses. Summary findings, based on the 85 responses from intermediaries**,** are presented in Section 2 of this report.

- Qualitative research undertaken with a separate sample of 15 intermediaries, which included eleven qualitative telephone interviews and an online survey of intermediaries, which received six responses[1]. The sample was selected purposively, from intermediaries who are regular users of CH bulk data products, in order to identify and provide coverage of the most frequent and intensive users of CH data. The interviews explored their use of CH data and other data sources, the costs and revenues attributed to the use of CH data, and the availability of substitute data sources. The findings of the qualitative research are presented in Section 3.

[1] Two of the intermediaries participated in both the online survey and a telephone interview.

1.4 Structure of this report

The remainder of this report is structured as follows:

- Section 2 presents findings from the WTP survey for intermediaries and draws comparisons with the wider sample of users of CH search and data services;
- Section 3 summarises the findings from the qualitative research with some key intermediaries who use bulk data products; and
- Section 4 presents the conclusions of the research with intermediaries.

The report also includes the following annexes:

- Annex 1 provides the topic guide used for the qualitative interviews with intermediaries;
- Annex 2 presents the questionnaire that was used for the online survey of intermediaries.

2 Results of the willingness to pay survey: intermediaries

This section summarises the findings of the willingness to pay (WTP) survey for intermediaries and draws comparisons with the results for all users of CH data.

2.1 The WTP survey

As described above, Report 2 provides details of the WTP survey that was undertaken as part of this study. It discusses: the design and development of the survey (including the choice task and time savings calculator); the sampling strategy and user profiles; and an analysis of the survey findings across all user types.

The WTP survey received a total of 608 responses across the pilot and main surveys[2], including 85 responses from businesses who stated that one of the main benefits of CH data for their organisation was *"information/data that we include in the products and services we sell to our customers"*. These 85 respondents have been defined as intermediaries, since they use CH data as an input to their own commercial products and services.

2.2 Sample profile

This section draws comparisons between the 85 intermediaries and the overall sample of 608 users. It has also been compared, where possible, with the profile of 7,763 users of CH data who responded to a supplemental user profile survey (also undertaken as part of this study)[3] and UK business statistics from the ONS[4]. This is intended to provide some comparative context, rather than to judge the representativeness of the sample, since the supplemental user profile survey results are likely to be weighted towards higher frequency users.

[2] The pilot and main survey data have been pooled because there were minimal changes to the survey between the pilot and main survey phases.

[3] The supplemental user profile survey was administered via pop-up links on the CHS / 'Search the Register', CHD, and WebCHeck websites over a 12-week period (December 2018 – February 2019). Users of the search services were invited to complete a short (5 minute) survey that compiled information on the type of user (business, research, or general public) and corresponding profile information.

[4] Office for National Statistics. (2018). Statistical Bulletin: UK business; activity, size and location: https://www.ons.gov.uk/businessindustryandtrade/business/activitysizeandlocation/datasets/ukbusinessactivitysizeandlocation

Respondent profile

The job title/role of the respondents is provided in Table 2.1. It shows that two in three (65%) of the 'intermediary' respondents had a professional occupation. This is higher than the overall sample of the WTP survey (44%) and the sample from the user profile survey (47%). Most of the other 'intermediary' respondents were either company directors (20%) or managers (6%).

Most respondents stated that they were the main user of CH data within their organisation. The findings suggest that the main users of CH data within intermediaries were also more likely to have a professional occupation compared to the other samples. This is consistent with the expected profile of intermediaries, in terms of using CH data in a professional capacity to create commercial products and services.

Table 2.1: Respondent job title/role

	Sample of intermediaries (n=85)	Total WTP survey (n=608)	User profile survey (n=5,491)
Company director	20%	30%	20%
Manager	6%	8%	9%
Professional occupation	65%	44%	47%
Technical occupation	0%	2%	2%
Administrative or secretarial	2%	7%	13%
Sales or customer services	2%	2%	4%
Other	5%	7%	6%
Total	100%	100%	100%

Sources: Direct user survey; Supplemental user profile survey

Organisation profile

The geographical distribution of the sample of intermediaries was broadly comparable with the samples from the total WTP survey and the user profile survey, and ONS data on the regional distribution of UK businesses. In all cases, between 63% and 68% of businesses were from the South East, London, South West and North West regions.

The profile of intermediaries was also similar to the other samples in terms of the type of organisation (Table 2.2). Most intermediaries in the sample were limited companies (69%), while 15% were partnerships, which is consistent with the overall profile of users of CH data and the population of UK businesses. However, there were slight differences in terms of a relatively high proportion of sole traders (13%) and an absence of non-profit or mutual organisations (0%) among the sample of intermediaries, which is also consistent with the expected profile of those using CH data to create commercial products and services.

Table 2.2: Type of organisation

		n	%
Limited company	Intermediaries	59	69%
	Total direct user sample	435	72%
	User profile survey		64%
	National profile		71%
Partnership	Intermediaries	13	15%
	Total direct user sample	51	8%
	User profile survey		15%
	National profile		17%
Sole trader	Intermediaries	11	13%
	Total direct user sample	54	9%
	User profile survey		5%
	National profile		8%
Public corporation	Intermediaries	2	2%
	Total direct user sample	18	3%
	User profile survey		5%
	National profile		0%
Central government	Intermediaries	0	0%
	Total direct user sample	5	1%
	User profile survey		2%
	National profile		0%
Local authority	Intermediaries	0	0%
	Total direct user sample	11	2%
	User profile survey		4%
	National profile		0%
Non-profit organisation or mutual (membership) organisation	Intermediaries	0	0%
	Total direct user sample	34	6%
	User profile survey		3%
	National profile		3%
Total			**100%**

Sources: Direct user survey; Supplemental user profile survey; ONS statistics.

Note: Intermediaries (n=85); Total direct user sample (n=608); User profile survey (n=5,477).

Table 2.3 presents the main activity of respondent organisations by sector. It shows that most of the intermediaries' activities were in the 'business administration and support services' (32%), 'professional, scientific and technical' (25%), 'finance and insurance' (21%) and 'information and communication' (11%) sectors. This concentration of activities is relatively high for each of these sectors, compared to the national structure of business activities. It is also relatively high compared to the total WTP survey and user profile survey samples in terms of 'business administration and support services' and 'information and communication' activities. Again, this is consistent with expectations of intermediaries, although one might have expected to see an even higher concentration of businesses within the 'information and communication' sector.

Table 2.3: Main activity of organisation

		n	%
Agriculture, forestry & fishing	Intermediaries	0	0%
	Total direct user sample	9	1%
	User profile survey		1%
	National profile		6%
Manufacturing	Intermediaries	1	1%
	Total direct user sample	22	4%
	User profile survey		5%
	National profile		6%
Construction	Intermediaries	0	0%
	Total direct user sample	18	3%
	User profile survey		5%
	National profile		12%
Motor trades	Intermediaries	1	1%
	Total direct user sample	5	1%
	User profile survey		1%
	National profile		3%
Wholesale	Intermediaries	1	1%
	Total direct user sample	9	1%
	User profile survey		2%
	National profile		4%
Retail	Intermediaries	0	0%
	Total direct user sample	11	2%
	User profile survey		2%
	National profile		8%
Transport & storage (warehousing)	Intermediaries	0	0%
	Total direct user sample	6	1%
	User profile survey		2%
	National profile		4%
Accommodation & food services	Intermediaries	1	1%
	Total direct user sample	5	1%
	User profile survey		1%
	National profile		6%
Information & communication	Intermediaries	9	11%
	Total direct user sample	39	6%
	User profile survey		7%
	National profile		8%
Finance & insurance	Intermediaries	18	21%
	Total direct user sample	119	20%
	User profile survey		24%
	National profile		2%
Property	Intermediaries	3	4%
	Total direct user sample	36	6%
	User profile survey		5%
	National profile		4%
Professional, scientific & technical	Intermediaries	21	25%
	Total direct user sample	157	26%
	User profile survey		22%
	National profile		18%
Business administration & support services	Intermediaries	27	32%
	Total direct user sample	100	16%
	User profile survey		11%
	National profile		8%
Public administration & defence	Intermediaries	1	1%
	Total direct user sample	15	2%
	User profile survey		4%
	National profile		0%
Education	Intermediaries	0	0%
	Total direct user sample	15	2%
	User profile survey		2%
	National profile		2%
Health	Intermediaries	0	0%
	Total direct user sample	11	2%
	User profile survey		2%
	National profile		4%
Arts, entertainment, recreation & other services	Intermediaries	2	2%
	Total direct user sample	31	5%
	User profile survey		3%
	National profile		6%
Total			**100%**

Sources: Direct user survey; Supplemental user profile survey; ONS statistics.

Note: Intermediaries (n=85); Total direct user sample (n=608); User profile survey (n=5,491).

Table 2.4 shows that the sample of intermediaries was generally positioned between the wider profiles of CH data users and the total UK business population in terms of organisation size. For example, it included a higher proportion of micro businesses (49%) than the wider user samples (43% and 22%) but a smaller proportion than the total UK business population (78%). Conversely, it included a smaller proportion of large businesses (11%) than the wider user samples (14% and 30%) but a higher proportion than the total UK business population (less than 1%).

Table 2.4: Number of employees

		n	%
0 - 4	Intermediaries	42	49%
	Total direct user sample	262	43%
	User profile survey		22%
	National profile		78%
5 - 9	Intermediaries	9	11%
	Total direct user sample	78	13%
	User profile survey		8%
	National profile		11%
10 - 19	Intermediaries	11	13%
	Total direct user sample	56	9%
	User profile survey		9%
	National profile		6%
20 - 49	Intermediaries	8	9%
	Total direct user sample	61	10%
	User profile survey		12%
	National profile		3%
50 - 99	Intermediaries	3	4%
	Total direct user sample	30	5%
	User profile survey		8%
	National profile		1%
100 - 249	Intermediaries	3	4%
	Total direct user sample	33	5%
	User profile survey		10%
	National profile		1%
250+	Intermediaries	9	11%
	Total direct user sample	88	14%
	User profile survey		30%
	National profile		0%
Total			**100%**

Sources: Direct user survey; Supplemental user profile survey; ONS statistics.

Note: Intermediaries (n=85); Total direct user sample (n=608); User profile survey (n=5,491).

The same was true of the sample of intermediaries in terms of annual turnover. Approximately half (49%) of the intermediaries had a turnover of less than £250,000, which was higher than the wider user samples (45% and 26%) but lower than the total UK business population (71%). Conversely, the intermediaries included a smaller proportion of businesses with a turnover in excess of £50m (13%) than the wider user samples (14% and 24%) but a higher proportion than the total UK business population (less than 1%).

Table 2.5: Annual turnover

		n	%
Up to £49,999	Intermediaries	19	22%
	Total direct user sample	123	20%
	User profile survey		11%
	National profile		16%
£50,000 - £99,999	Intermediaries	12	14%
	Total direct user sample	79	13%
	User profile survey		7%
	National profile		23%
£100,000 - £249,999	Intermediaries	11	13%
	Total direct user sample	70	12%
	User profile survey		8%
	National profile		32%
£250,000 - £499,999	Intermediaries	6	7%
	Total direct user sample	44	7%
	User profile survey		5%
	National profile		12%
£500,000 - £999,999	Intermediaries	10	12%
	Total direct user sample	39	6%
	User profile survey		7%
	National profile		7%
£1,000,000 - £1,999,999	Intermediaries	8	9%
	Total direct user sample	49	8%
	User profile survey		9%
	National profile		4%
£2,000,000 - £4,999,999	Intermediaries	4	5%
	Total direct user sample	46	8%
	User profile survey		9%
	National profile		3%
£5,000,000 - £9,999,999	Intermediaries	3	4%
	Total direct user sample	26	4%
	User profile survey		7%
	National profile		1%
£10,000,000 - £49,999,999	Intermediaries	1	1%
	Total direct user sample	47	8%
	User profile survey		13%
	National profile		1%
£50,000,000 or more	Intermediaries	11	13%
	Total direct user sample	85	14%
	User profile survey		24%
	National profile		0%
Total			**100%**

Sources: Direct user survey; Supplemental user profile survey; ONS statistics.

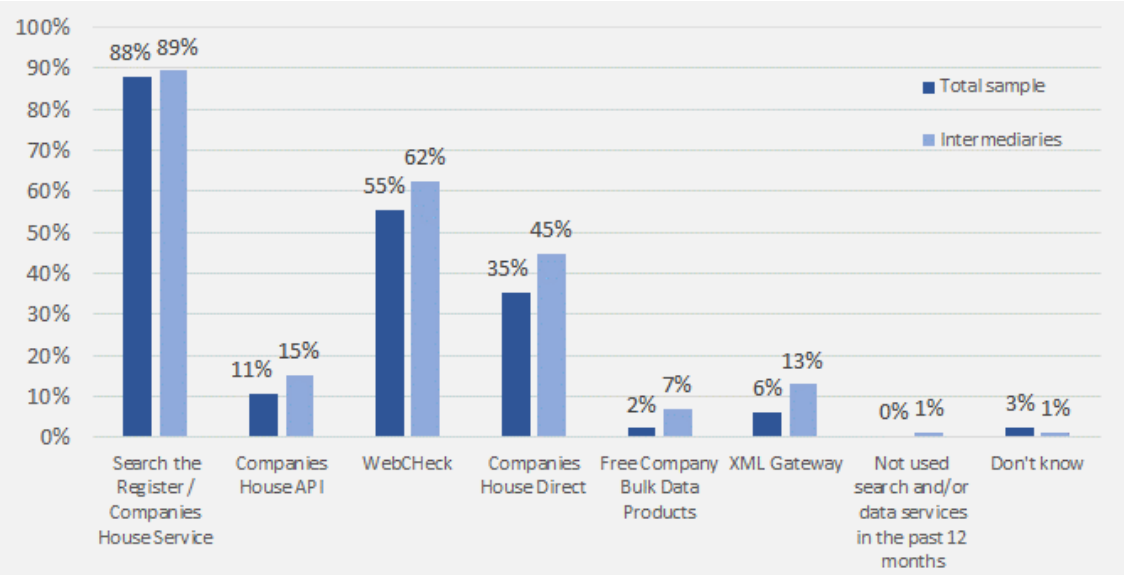
Note: Intermediaries (n=85); Total direct user sample (n=608); User profile survey (n=5,491).

2.2 Use of CH data

The following figures summarise the use of CH data services among the sample, showing the total levels of usage (Figure 2.1) and the services used most often (Figure 2.2). As expected, the sample of intermediaries reported relatively high levels of use of CH data. All except one had used CH data services during the last year, and all had used them within the last two years.

Figure 2.1 shows that Companies House Service (CHS) / 'Search the register' was the most common service among the intermediaries (used by 89%) and the total WTP survey sample (used by 88%). The sample of intermediaries also reported relatively high usage of WebCHeck (used by 62%), Companies House Direct (45%), Companies House API (15%), XML Gateway (13%) and free bulk data products (7%).

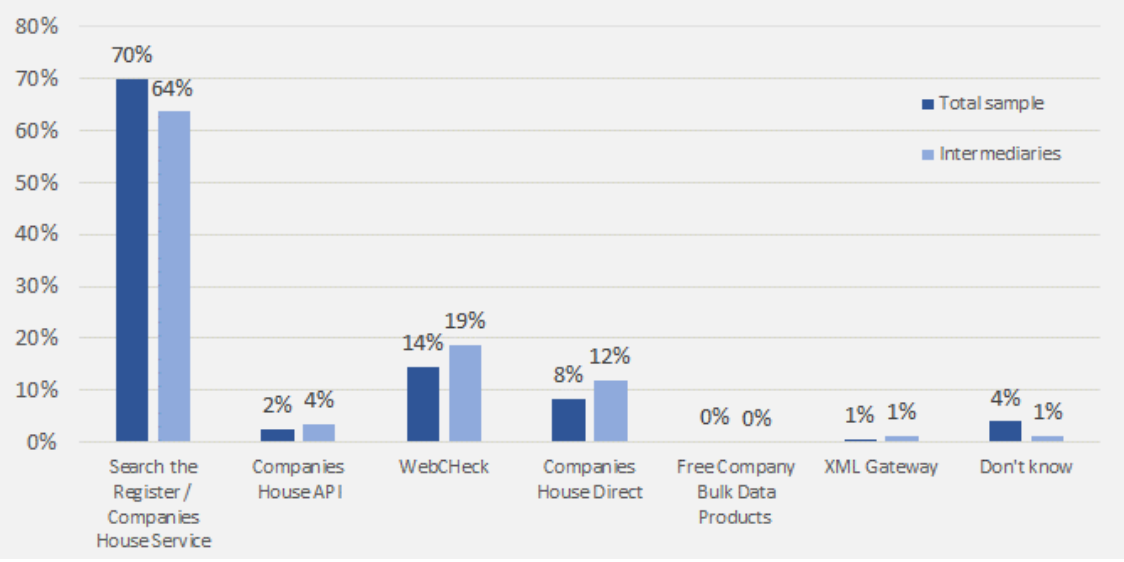
Figure 2.1: CH search and data services used in the past 12 months



Source: Direct user survey. Note: Intermediaries (n=85); Total direct user sample (n=608).

Figure 2.2 shows that the main services used by the sample of intermediaries were similar to the total sample. CHS / 'Search the register' was reported to be the main service used by intermediaries (for 64% of the sample), followed by WebCHeck (19%), Companies House Direct (12%) and Companies House API (4%). None of the intermediaries listed the free bulk data products as the service used most often. This is a clear difference from the firms interviewed in the qualitative research, which tended to use bulk data most frequently. A comparison of samples found no duplication between respondents to the WTP survey and the qualitative interviews. This suggests that intermediaries are not a homogeneous group and have different roles in the market.

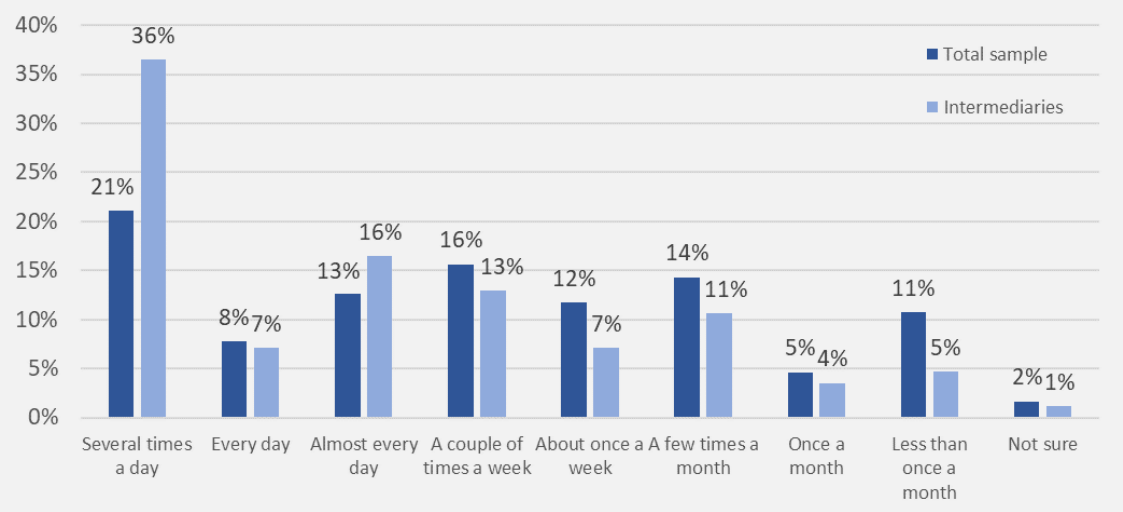
Figure 2.2: CH search and data services used most often



Source: Direct user survey. Note: Intermediaries (n=85); Total direct user sample (n=608).

Figure 2.3 summarises respondents' reported frequency of use of CH data services. It suggests that these intermediaries are relatively frequent users of CH data, which is consistent with expectations for this group of users. For example, it shows that 36% of the intermediaries reported using CH data several times per day, compared to 21% of the total WTP survey sample (and 33% of the supplemental user profile survey, although this information has not been included).

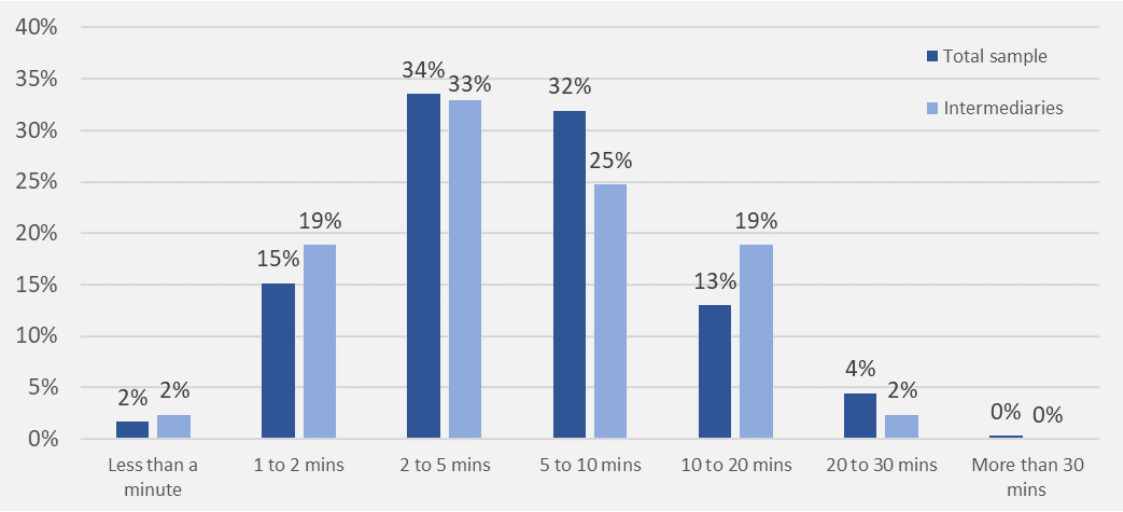
Figure 2.3: Frequency of use



Source: Direct user survey. Note: Intermediaries (n=85); Total direct user sample (n=608).

The average time spent using CH data was reported to be similar for the sample of intermediaries and the total WTP sample. Most intermediaries (58%) spend between two and ten minutes for each use of CH data. However, the sample of intermediaries also included a slightly higher proportion of users spending less than two minutes (21%) and more than ten minutes (21%) for each use, relative to the total sample.

Figure 2.4: Average time for each use of CH search or data services



Source: Direct user survey. Note: Intermediaries (n=85); Total direct user sample (n=608).

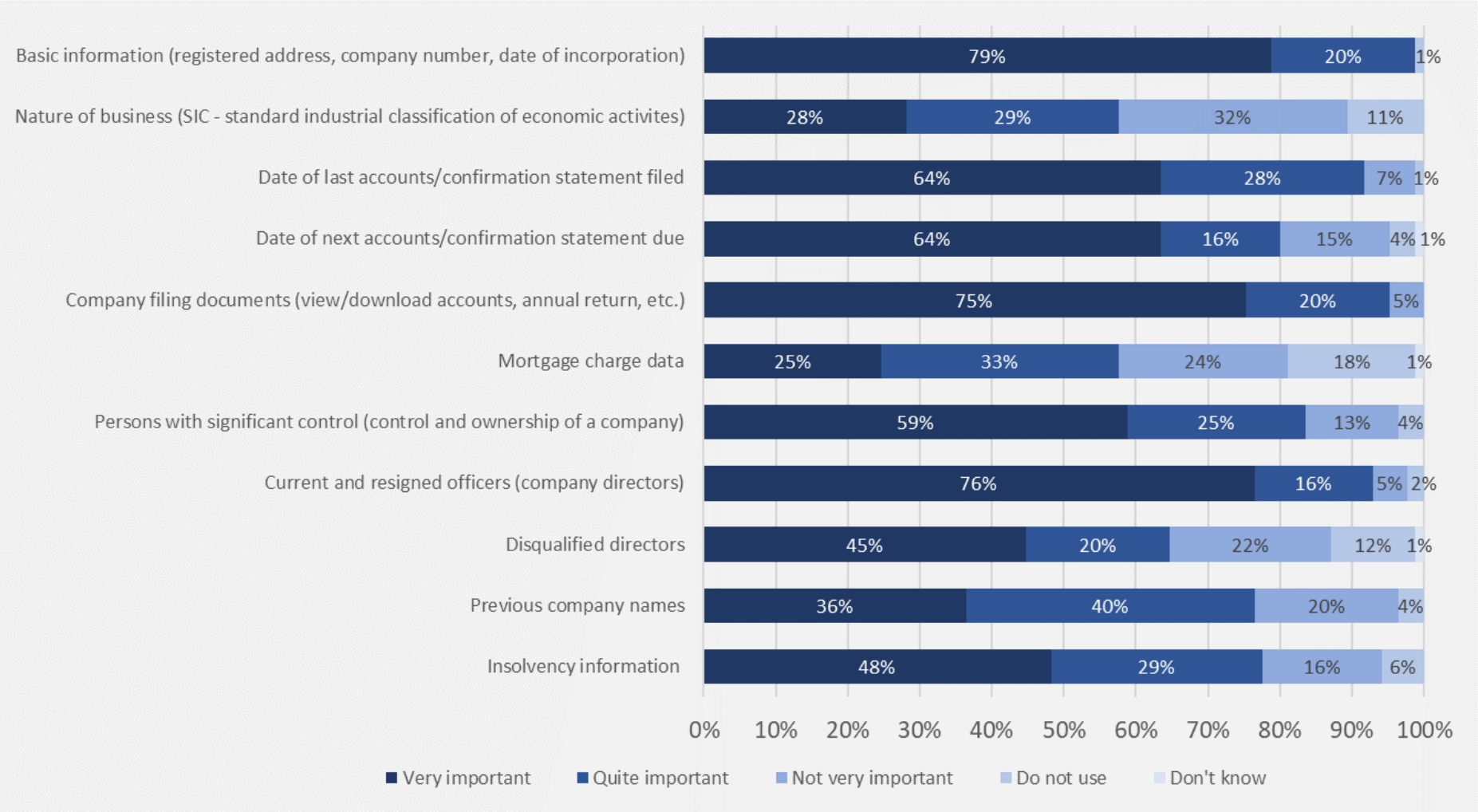
Uses of CH information and data

Respondents to the WTP survey were asked to rate the importance of the different aspects of the CH information and data to their organisation. Figure 2.5 shows that most of the intermediaries (>50%) indicated that all aspects of the information were ‘very important’ or ‘quite important’ for their organisation.

The most important pieces of data for the intermediaries were reported to be the basic company information (company number, registered address, date of incorporation, etc.), company filing documents (accounts, annual returns, etc.) and the current and resigned officers (company directors). At least 75% of the sample of intermediaries stated that this information was 'very important' for their organisation.

Other important aspects of the data included: the dates of filing the last and next accounts / confirmation statements; 'persons with significant control' (PSC) data; insolvency data, and data on previous company names. At least 75% of the intermediaries reported that each of these aspects was 'very important' or 'quite important' for their organisation.

Figure 2.5: Importance of different aspects of CH services and data for intermediaries



Source: Direct user survey. Note: Intermediaries (n=85); Total direct user sample (n=608).

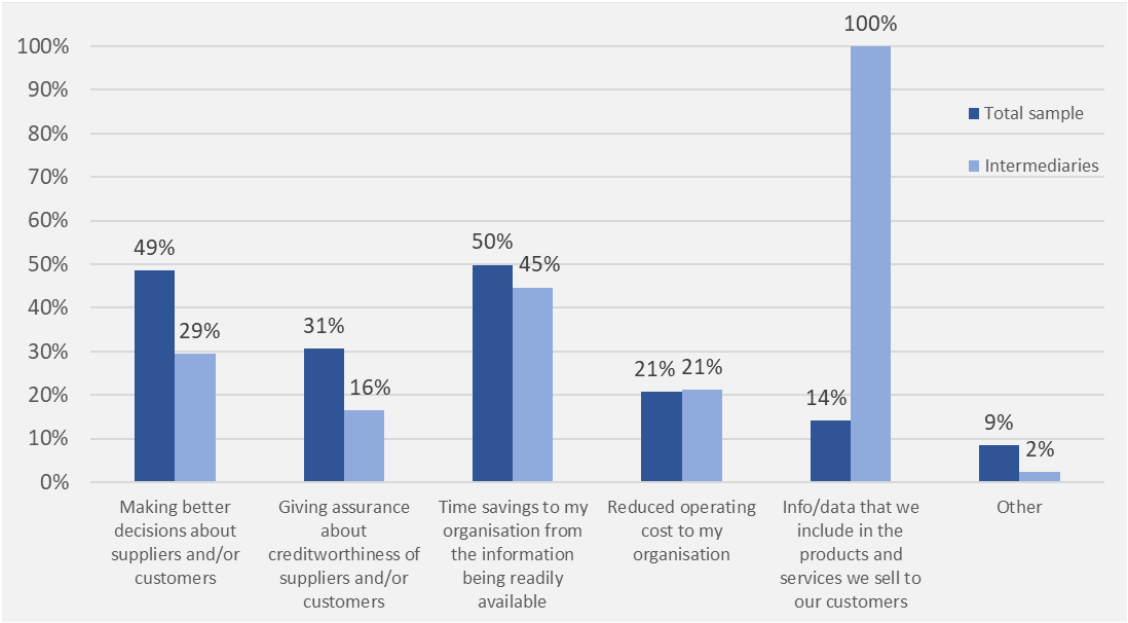
2.4 Benefits of CH data

The WTP survey also asked respondents about the benefits that CH data had provided for their organisations. Figure 2.6 shows all of the benefits reported by respondents, while Figure 2.7 focuses on the main benefit for each organisation.

Clearly the figures show that all intermediaries reported benefiting from including CH data in the products and services they sell to customers, while two in three (67%) also stated that this was the main benefit of CH data for their organisation. This was significantly higher than the total WTP sample, for which 14% reported benefiting in this way, with 10% suggesting that this was the main benefit of CH data.

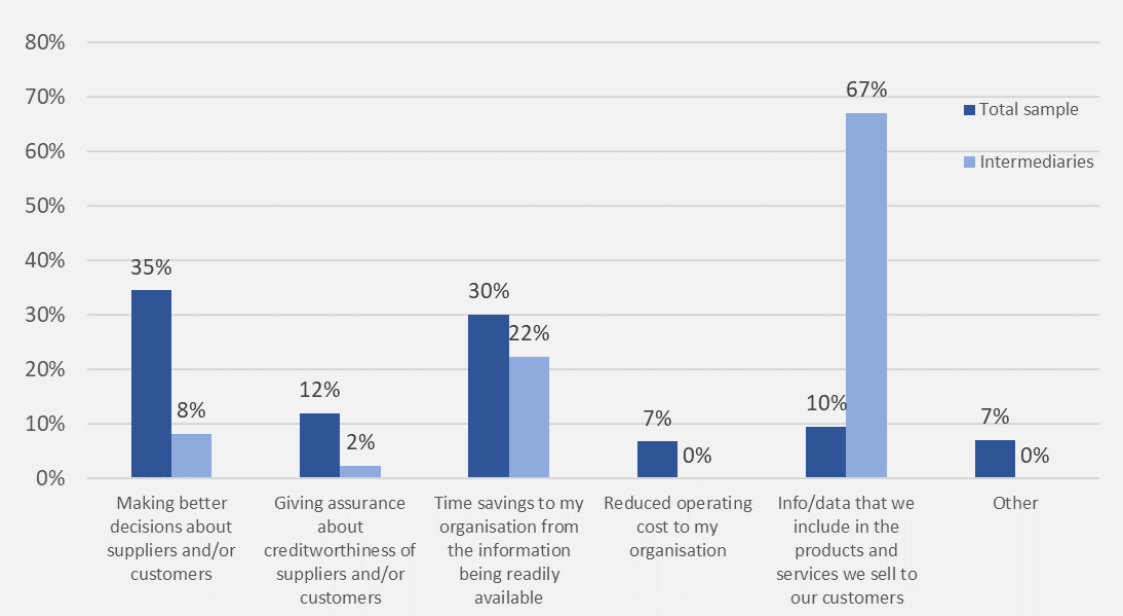
The next most common benefit for intermediaries was the time savings provided by CH data. This was a benefit for almost half (45%) of the sample of intermediaries and was the main benefit for 22%. A smaller proportion of the intermediaries also reported benefiting from reduced operating costs and being able to make better decisions and receiving assurance about suppliers and customers.

Figure 2.6: All beneficial outcomes of CH information and data



Source: Direct user survey. Note: Intermediaries (n=85); Total direct user sample (n=608).

Figure 2.7: Main benefit of CH information and data



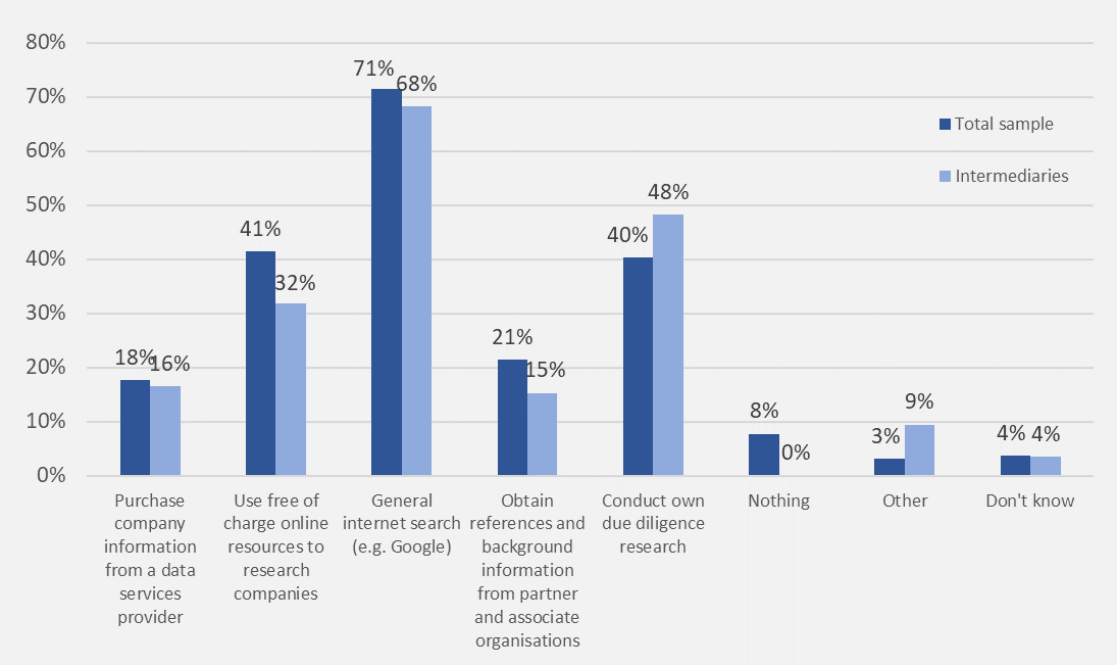
Source: Direct user survey. Note: Intermediaries (n=85); Total direct user sample (n=608).

Overall levels of satisfaction with CH data were high among the sample of intermediaries, with 62% stating that they were very satisfied and 24% stating they were fairly satisfied with CH data services (which was broadly similar to the equivalent figures of 67% and 21% for the total WTP survey sample). Confidence in the accuracy and reliability of CH data was also reported to be high, with 85% of both the intermediaries and the total sample stating that they were very or extremely confident in the accuracy and reliability of the data.

2.5 Use of alternative services and products

The WTP survey also asked respondents to indicate the alternative products and services their organisations used to source company information and data. Figure 2.8 shows that the most common alternative sources for the sample of intermediaries were to use general internet searches (used by 68% of the sample), to conduct their own research (48%) and use other free online sources (32%). The total WTP survey sample reported similar results, although intermediaries were slightly more likely to conduct their own research and less likely to use free online resources.

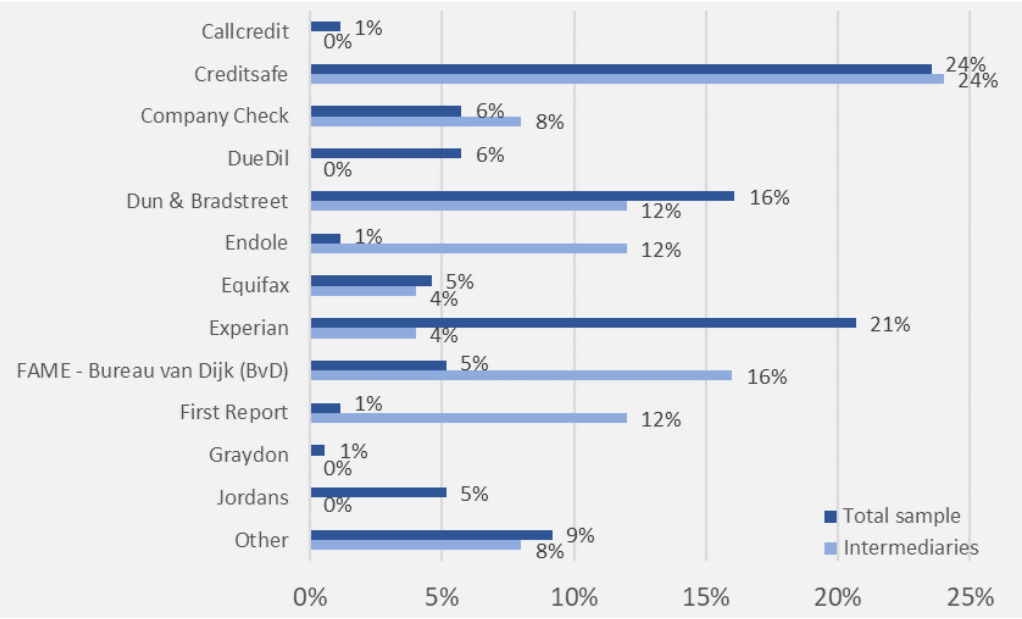
Figure 2.8: Use of alternative data and information sources



Source: Direct user survey. Note: Intermediaries (n=85); Total direct user sample (n=608).

Fourteen of the intermediaries (16% of the sample) reported purchasing or subscribing to commercial data services. Figure 2.9 shows that these intermediaries reported using a range of different services including Creditsafe, FAME/Bureau van Dijk, Dun & Bradstreet, Endole, First Report, Company Check, Equifax and Experian. The findings suggest that the sample of intermediaries was more likely than the total sample to use FAME, Endole and First Report, and less likely to use services from Experian and Dun and Bradstreet.

Figure 2.9: Data services respondents' organisation purchases or subscribes to



Source: Direct user survey. Note: Intermediaries (n=25); Total direct user sample (n=107).

2.6 Annual user benefit estimates

Intermediaries' WTP[5] for specific types of information and data currently provided by CH has been calculated from the choice model estimation results reported in Report 2[6]. The results are summarised below in Table 2.6 in terms of mean (average) unit value values. They represent the annual benefit to intermediaries from the provision of each type of information (basic information about a company; PSC data; and annual reports and statements). The total value of user benefits for intermediaries – summed across all attributes – is approximately £2,000 per year per user, which is consistent with the overall estimates for all direct users, presented in Report 2, and has a reasonable uncertainty range given the sample size (85 respondents). The lower/upper bound range is approximately £1,200/£2,900, based on the corresponding 95% confidence interval estimates. As expected, the interval estimate is wider for intermediaries, which is a result of the smaller sample size and lower level of precision for these results compared to the overall sample.

Table 2.6: User benefits for intermediaries – mean (average) WTP per year (£/year/user)

Company information and data attribute	Central	Lower	Upper
Company information (basic details)	960.54	541.05	1,380.03
PSC data	0	0	0
Annual reports & financial statements	1,066.07	642.16	1,489.98
Total	2,026.61	1,183.21	2,870.01

Notes: WTP calculated from non-linear model (CL dummy-coded model; pooled sample – Report 2; Annex 7; Section 3.7). Lower – upper bounds are 95% confidence interval. All values statistically significant at the 1% level.

The results suggest that intermediaries attribute the greatest proportion of the overall annual benefit value to the provision of financial information (e.g. annual reports and financial statements). This represents approximately half (52%) of the total benefit (approximately £1,100 per year). Company information accounts for just under half (48%) of the total benefit (approximately £950 per year).

Based on these results, no additional value is associated with the provision of PSC data. This, however, likely reflects the limitations of the sample size in terms of the precision of individual parameter estimates and the underlying model assumptions. Validity testing analysis indicates a positive value for PSC data based on alternative assumptions for the

[5] WTP is calculated as the ratio of estimated coefficients for the attribute (level) and the annual cost (marginal utility of money) parameter; i.e. WTP = $-\beta x / \beta cost$, where x is the attribute level and the β's are the coefficient values.

[6] BEIS (2019) Valuing the user benefits of Companies House data: Report 2: Direct Users of Companies House Data [Section 6; Annex 7]

distribution of parameter estimates (see below), while the qualitative research (Section 3) also highlights the perceived importance of PSC data among intermediaries.

An equivalent value per use (i.e. visit to/use of CH search service) can be obtained by dividing the annual benefit estimate by the average number of times that CH search services are accessed per year (262 times per intermediary user). This gives a benefit value of approximately £8 per intermediary user per individual use (with a lower/upper bound of approximately £5/£11).

Validity assessment

The validity of the annual user benefit values is assessed through two supplemental analyses: (i) a comparative assessment that uses an alternative econometric specification to estimate user WTP using alternative assumptions concerning the distribution of parameter estimates (see Report 2; Annex 7; WTP-space estimation); and (ii) comparison to results from the time saving cost calculator (see Report 2; Section 3.6; Annex 6), which conceptually provides a lower-bound resource cost-based estimate of user benefits.

Figure 2.10 compares the estimated user benefits (presented above in Table 2.6) to the results from the alternative WTP-space estimation. Overall, there is consistency in the two sets of values based on the overlapping estimates, using 95% confidence intervals. This means that it is not possible to conclude that the valuations are significantly different from each other. Overall the total value summed for the WTP-space estimation is approximately £1,800 per intermediary user per year, compared to approximately £2,000 for the results presented above in Table 2.6. Notwithstanding these findings, a notable result is the positive value for PSC data (approximately £46 per intermediary user per year). In practice this demonstrates the uncertainty margins for the analysis of the intermediaries sub-sample and shows how results from small data samples can be sensitive to assumptions concerning the distribution of parameter estimates.

Figure 2.10: Comparison of utility space and WTP space benefit estimates – mean (average) and interval estimates for WTP per year for intermediaries (£/year/user)

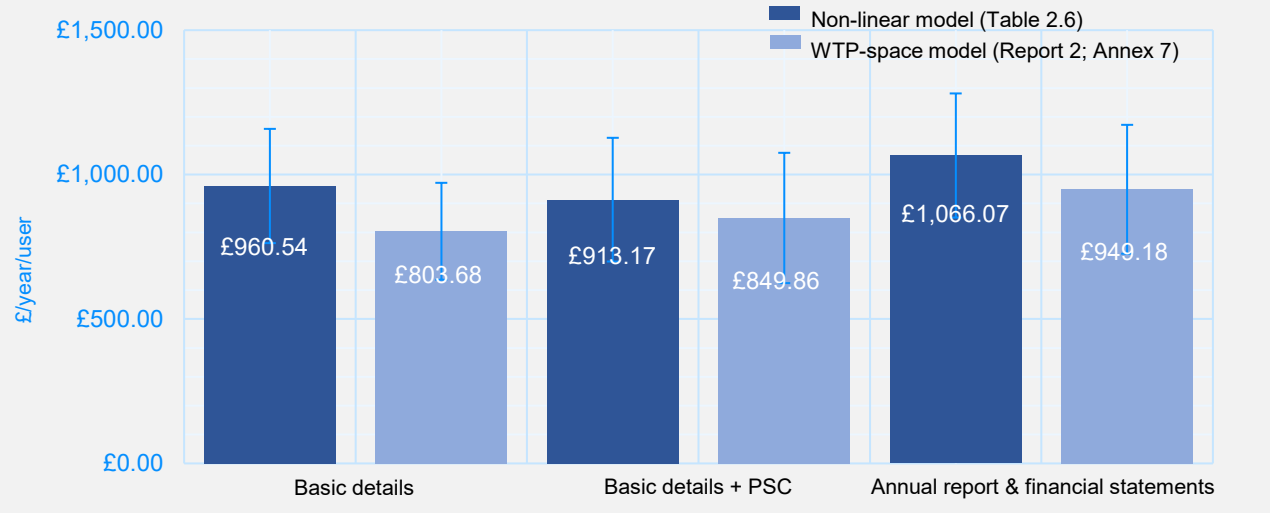


Table 2.7 summarises results from the time savings calculator questions for intermediaries. As described in Report 2 (Section 3.6) these can be interpreted as a proxy for the value of company information and data to users, representing a lower-bound cost-based comparator or benchmark that can help assess the plausibility of the user WTP estimates. Result are reported per user per year, and are weighted according to the use of company search services by the job title/role (at the sample average). The value of time savings is calculated based on ONS Annual Survey of Hours and Earnings (ASHE) statistics for gross hourly pay[7].

Table 2.7: Estimated timing savings (£/year/user)

	Time saving per year (hh:mm:ss)	Estimated value (gross hourly pay)			
		Mean	Median	10th percentile	90th percentile
Company director	9:35:32	249.87	197.02	91.89	464.74
Manager	3:28:41	90.60	71.44	33.32	168.51
Professional occupations	11:22:04	252.93	227.70	143.80	382.98
Technical occupations	0:31:51	9.49	8.14	5.18	14.19
Administrative or secretarial	3:46:15	46.91	40.46	29.87	70.03
Sales or customer service	0:54:55	9.10	7.74	6.84	12.32
Other	1:06:38	11.04	9.40	8.30	14.95
Total	**30:45:56**	**669.95**	**561.90**	**319.19**	**1,127.71**
Total incl. non-wage labour costs*	**-**	**815.87**	**684.28**	**388.71**	**1373.33**

Notes: * Uplift for non-wage labour costs of 21.78% to cover additional employee benefits such as pensions, National Insurance contributions, sickness pay as well as maternity and paternity pay. The uplift factor is consistent with the approach applied in Impact Assessments by BEIS (Pers. Comm. BEIS, July 2019). Source: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lc_lci_lev&lang=en

The results in Table 2.7 show that the average time saving for intermediaries is approximately 31 hours per year, with the benefits primarily accruing in terms of the time of employees in professional occupations, such as accountants (37%), and the time of company directors (31%). Based on average wage rates for these occupations, this corresponds to a weighted average value of around £600 to £700 per intermediary user per year (based on the median and mean gross pay, respectively). Factoring in other costs to employers (e.g. National Insurance, pensions, etc.), the uplifted values are approximately £700 - £800 per user per year. Applying the wider distribution of wage rates (using the 10th and 90th percentiles for illustrative purposes) gives a range of approximately £400 to £1,400 per intermediary user per year.

[7] See: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/allemployeesashetable1

Overall there is an encouraging degree of consistency between the time saving estimates and user WTP estimates. As expected, the former estimates are lower than the calculated benefit values for users, but are of a similar order of magnitude. This represents a form of convergent validity, demonstrating that the estimated WTP values can be reconciled against an alternative measure of benefits (based on avoided costs). By and large this should be interpreted as a form of ‘validity by association’, with the findings reflecting reasonable prior expectations as to why the comparative results would differ within reasonable bounds.

3 Results of the qualitative research: intermediaries

This section presents the results of qualitative research undertaken with a sample of key intermediaries who use CH bulk data products as an input to their own commercial products and services.

3.1 Intermediaries who also use CH bulk data

A subset of intermediaries can also be identified through their use of 'bulk' data products. Bulk data products provide users with large quantities of CH data that intermediaries can use as an input to their own commercial products and services. These users of bulk data services are therefore likely to include the most frequent and significant users of CH data, and those for whom CH data generates the largest revenues.

The ‘bulk’ data products provided by CH include a broad range of information such as:

- daily data files including new incorporations, company appointments, directory updates, liquidations, mortgages, accounts data, and some bespoke data products developed for individual users;
- weekly data files including the weekly gazette and information on disqualified directors;
- monthly data files on DVD ROM and some bespoke monthly data products covering dissolutions and liquidations; and
- additional bespoke data files prepared on quarterly basis including data on liquidations.

While most CH data can be accessed anonymously, users of bulk data must register in order to access these products. CH provided anonymised information on all 132 of the intermediaries that were accessing bulk data products from CH at the time of the research. This provided some basic information on the characteristics of these users and their transactions of bulk data products, which are described in more detail below.

Qualitative research was also undertaken with 15 of these intermediaries, who were purposively selected to identify a sample of significant users of CH data. While this sample represents 11% of the total number of intermediaries that are currently accessing bulk data products (132), it is also expected to include a relatively large proportion of the benefit of CH data to this sub-group of users as it includes the main credit reference agencies (including all seven members of the Business Information Providers Association – BIPA).

Furthermore, the user information provided by CH showed that the research sample included relatively heavy users of bulk data products who accounted for 21% of the total number of bulk products accessed by the sample of intermediaries. Many of the other intermediaries, that were excluded from the research sample, are expected to be much smaller organisations, which is consistent with the findings of the WTP survey, presented in Section 2, that suggests that most intermediaries in that sample were micro or very small businesses.

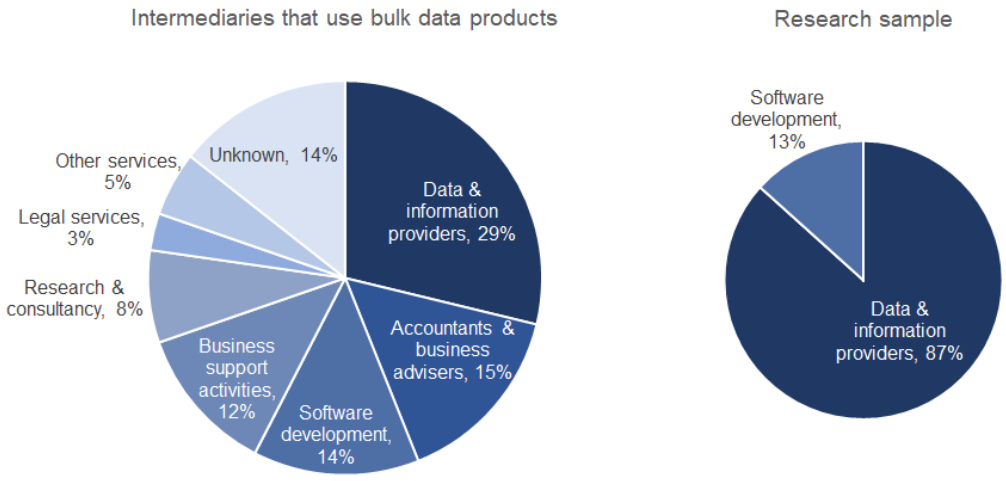
The characteristics of the interview sample are also described below, followed by a summary of the key findings of the qualitative interviews.

3.2 Characteristics of intermediaries that use CH bulk data

This section uses the anonymised information provided by CH to summarise the characteristics of the 132 intermediaries accessing bulk data and compares the characteristics of the sample of 15 intermediaries that were interviewed and/or responded to an online survey as part of this study. Figure 3.1 shows the sectors associated with this sample of intermediaries. While sector information was not available for 14% of users, the data suggest that the most common sector was 'data and information providers', accounting for at least 29% of the intermediaries that use bulk data products. Other common sectors included accountants and business advisers (at least 15% of the total), software development (at least 14%) and business support activities (at least 12%).

The research sample was more focused on data and information providers and software development. As described above, this is likely to be due to the research sample focusing on the higher value users within these sectors (such as credit reference agencies). Therefore, while the research sample is not representative of the wider population (of intermediaires that use bulk data products) in terms of the breadth of sector coverage, it is likely that it captures the activities that add most value to CH data amongst these users.

Figure 3.1: Type of organisation – intermediaries that use CH bulk data

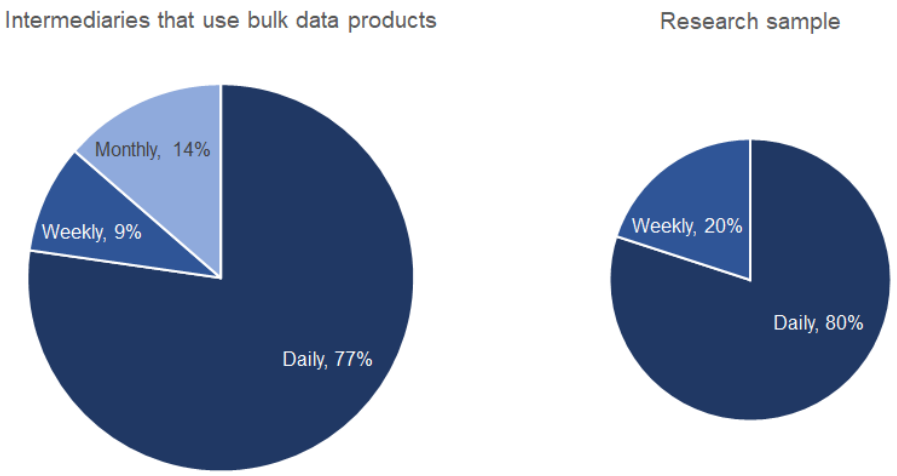


Source: ICF analysis of CH records for 132 intermediaries that use bulk data and a sub-sample of 15 intermediaries.

CH also provided information on the frequency of use of the bulk data products accessed by intermediaries, which is summarised in Figure 3.2. This confirms that intermediaries tend to be very frequent users of bulk data from CH, with more than three-quarters (77%) accessing the data on a daily basis. The remaining intermediaries are split between those accessing bulk data on a weekly (9%) or monthly (14%) basis.

The characteristics of the research sample are similar with twelve of the fifteen respondents (80%) accessing bulk data on a daily basis and the other three (20%) accessing data every week. This might suggest slightly more frequent usage of bulk data amongst the research sample, which is again likely to represent the more intensive usage amongst these users.

Figure 3.2: Frequency of access to 'bulk' data – intermediaries that use CH bulk data

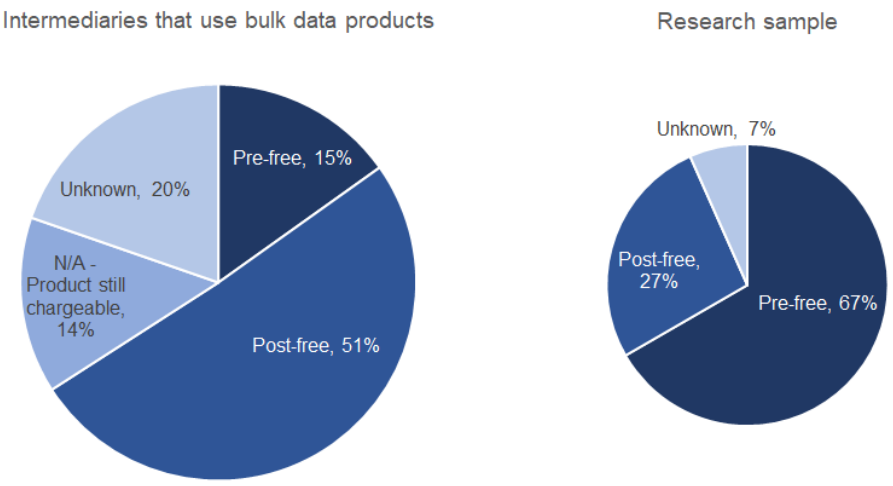


Source: ICF analysis of CH records for 132 intermediaries that use bulk data and a sub-sample of 15 intermediaries.

Figure 3.3 summarises CH data on the length of time that each intermediary user has been accessing bulk data from CH. This information has been consolidated into those that were accessing bulk data before and after it was made available free of charge. It shows that approximately half of these intermediaries (51%) have only been accessing bulk data products since they became available free of charge, compared to only 15% that used to pay a fee to access these products. For the other intermediaries, this information was not available in 20% of cases, while a further 14% are accessing bulk data products for which charges still apply (e.g. DVD ROM).

In contrast, ten of the fifteen intermediaries in the research sample (67%) had been accessing bulk data from CH for a longer period of time and used to pay fees to access the data. Only four intermediaries (27% of the research sample) had started accessing bulk data after it became available free of charge. This provides further indication that the research sample includes a concentration of higher value-added intermediaries, that have been using bulk data as an input to their commercial products over a relatively long period of time and were previously willing to pay fees to access the data.

Figure 3.3: Date of first download of bulk data – intermediaries that use CH bulk data

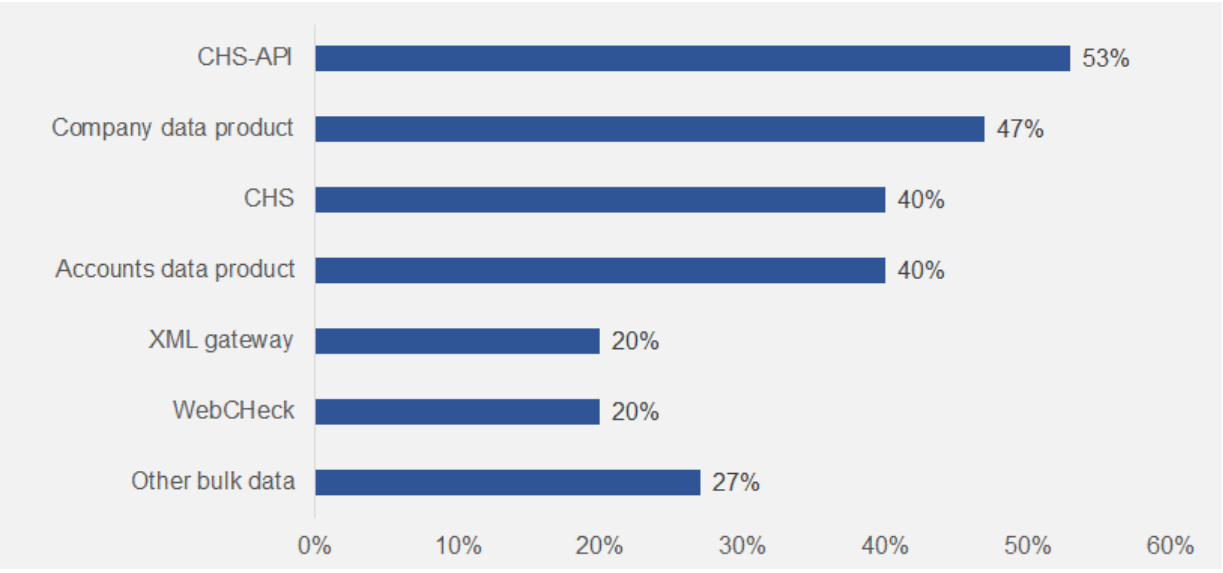


Source: ICF analysis of CH records for 132 intermediaries that use bulk data and a sub-sample of 15 intermediaries.

3.3 Use of CH data

The qualitative interviews and online survey collected information on the full range of CH data products and services that were being used by the sample of intermediaries. Figure 3.4 shows that the sample was using a range of different data products with no single product being used by a clear majority of users. The most commonly used products were the CHS-API service and the company data product. Each of these products was reported to be used by around half of the sample (i.e. by seven or eight of the fifteen users), while six of the intermediaries (40%) reported using the accounts data product and CHS service. A smaller number of intermediaries also reported using other bulk data, WebCHeck and the XML gateway. Approximately half of the sample (47%) also reported using PSC data.

Figure 3.4: Use of CH data products and service by the sample of intermediaries



Source: ICF analysis of 15 qualitative interviews with intermediaries that use bulk data products.

The sample of intermediaries was also asked which of the CH data products were most valuable for their businesses. Their responses closely matched the most commonly used products and suggested that:

- The CHS-API was the most valuable product for several of the intermediaries because it enables them to talk directly to the CH data and systems and also access data that are not available from other products. One user felt that the API service would offer even greater opportunities in the future due to its ability to provide more flexible, tailored access to the data.

- The bulk downloads provided by the company data product and accounts data product were the core products for several other intermediaries in this sample and perceived to provide the greatest value. These users valued the breadth of information available and the ability to search for companies, access financial information and check the filing history, which provides a useful timeline of a company including changes over time (e.g. addresses).

- The recent launch of PSC data had been particularly valuable for some of the intermediaries in terms of providing important information on company owners that can be difficult to access from other sources. One user suggested that the introduction of PSC data had enabled their organisation to enhance their products that focus on money-laundering.

Other intermediaries in the sample had suggested that all CH data products are complementary and equally valuable. However, one user reported that the mortgage data was probably the least valuable of the products that they access, although it is still useful in helping to build a broad and detailed picture of companies.

Changes in the use and value of CH data over time

Nearly all of the sample of intermediaries stated that their use of CH data had remained relatively constant over time. As stated above, most of the sample had been using CH data for a relatively long period of time, with ten of the fifteen respondents using bulk data before free access was introduced. Moreover, most of these respondents suggested that the introduction of free data had only had a minimal impact on their levels of usage of CH data, with only one or two users reporting access to additional datasets, such as data on liquidations, following the introduction of free data. Most intermediaries in the sample suggested that their usage of CH data had remained constant over time. For example, one user reported that it was nice to be able to access the data for free, as this had delivered cost savings of approximately £100,000 per year, but this had not affected their usage and they would have continued to access the same data if the charges had remained in place. To some extent, this is likely to be due to the focus of the sample on the most significant users that add the greatest value to CH data, for whom the previous data fees represented a relatively low proportion of the value that was generated by the data.

A few of the other intermediaries reported accessing more data over time but suggested that this was due to changes made by CH in terms of adding additional data fields, switching the frequency of datasets from weekly to daily delivery, or adding new products such as the introduction of PSC data.

PSC data

Most of the sample of intermediaries considered the introduction of PSC data to be the most significant recent development to CH data, since the introduction of free data. Some respondents felt it provided much needed data that enabled them to improve their products and services and meet the needs of their customers more effectively. For example, two users reported developing new products as a direct result of the introduction of PSC data. The PSC data had enabled them to develop new 'due diligence' products and offer money-laundering checks against the PSC data.

However, other users in the sample suggested that the introduction of PSC data had not had much of an impact. Some also expressed frustrations that the PSC data was not as comprehensive and transparent as they had expected due to changes to the confirmation statements, which removed the requirement for companies to provide a full list of all owners/shareholders.

Use of complementary data sources

The sample of intermediaries reported using a broad range of complementary data sources, in addition to CH data, when developing their products and services. Examples included:

- products and services of other intermediaries;
- financial information relating to bank accounts, balances, loan agreements, payments, etc.; and
- data from other sources including County Court Judgements (CCJs), insolvencies, company addresses, non-UK companies, profit warnings, secure creditors, and other proprietary data (such as data from the Registry Trust, Land Registry, Gazette, Post Office, Ministry of Justice, etc.).

In most cases the complementary data sources were being used to provide additional data that was not available from CH. For example, much of the complementary sources were providing data for individuals rather than businesses. The sample also reported that CH has a strong reputation for providing good quality and reliable data, although some respondents suggested that other sources can be more flexible and make changes more quickly to address issues that may arise. There were also examples where other sources can provide quicker access to data that is also provided by CH, such as data on insolvencies.

Most of the intermediaries in the sample also stated that it was relatively easy to combine CH data with data from other sources. In most cases, these users had developed systems to automate the matching process and many were using the company numbers in the CH data as their unique key, against which to match all other sources. Moreover, most examples of matching issues related to other sources, such as Registry Trust data, which do not have unique codes, and data on CCJs. For example, one of the intermediaries suggested that, on average, it is not possible to match around 30% of CCJs, so these are currently ignored. This user also reported trying to encourage the Ministry of Justice to use the CH company number in their data to support more effective matching of CCJs to companies.

3.4 Estimating the value of CH data for intermediaries

The qualitative interviews with intermediaries also collected information on their costs of accessing and processing CH data and the revenues generated from products and services that use CH data. They also sought to attribute costs and revenues to the use of CH data. This section presents these findings for this sample of intermediaries that use bulk data products.

The total turnover of this sample of intermediaries was £921m. This equates to an average turnover across the sample of approximately £60m per user, although this figure has been inflated by several very large businesses with revenues in excess of £100m, as the sample also included several micro businesses with revenues of less than £1m. For the purposes of this analysis, the sample has been split into two sub-groups:

- One group includes the largest eight intermediaries within the sample, including the seven members of BIPA. These eight organisations have a combined UK turnover of more than £900m, accounting for 99% of the total turnover of the sample. They are referred to as 'larger intermediaries' in the following analysis.

- The second group includes the other seven intermediaries in the sample. These seven organisations are considerably smaller and have a combined turnover of approximately £10m, representing just 1% of the total across the sample.However, these intermediaries are still expected to be larger than the intermediaries identified through the WTP survey, and are referred to as 'mid-sized intermediaries' in the following analysis.

Estimates of total costs have also been calculated for each of the intermediaries in the sample. This uses data from the Annual Business Survey (ABS)[8] to calculate total costs as

[8] ONS (November 2018) UK non-financial business economy (Annual business survey): sections A to S – 2017 provisional results. Available at: https://www.ons.gov.uk/file?uri=/businessindustryandtrade/business/businessservices/datasets/uknonfinancialbusinesseconomyannualbusinesssurveysectionsas/current/abssectionsas.xls

a percentage of turnover within their respective sector. The combined figures across the sample of intermediaries estimate their combined costs to total almost £700m[9] (i.e. 76% of their turnover respectively). Separate figures are provided for the two sub-groups, which show that the eight larger intermediaries account for around 99% of the turnover and costs of the sample as a whole.

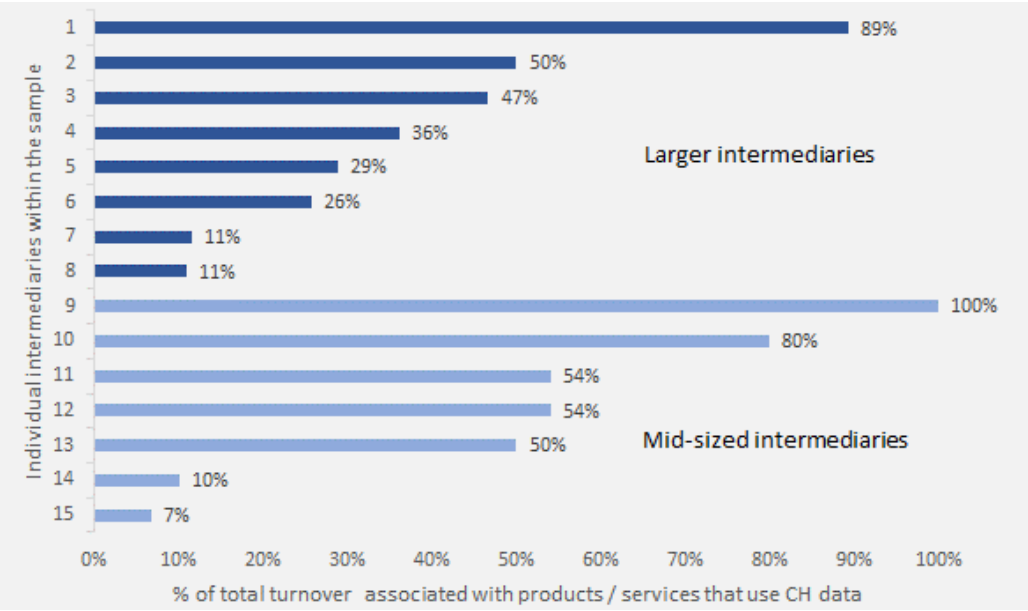
The interviews and online survey also asked the intermediaries how much of their overall business turnover related to products and services that use CH data as an input. There was significant variance in the estimates across the sample, as shown in Figure 3.5, for example:

- the group of larger intermediaries suggested that products and services that use CH data account for between 11% and 89% of their total turnover, although most of the sub-sample reported proportions of between 25% and 50%.

- the group of mid-sized intermediaries provided an even larger range of estimates, varying from 7% to 100% of their overall turnover, although most of this sub-sample reported proportions of at least 50%.

Overall, the group of mid-sized intermediaries suggested that these products represented a higher proportion of their overall turnover. This was also the case within each of the sub-groups, with smaller organisations typically reporting a higher proportion than the larger businesses that tended to provide a greater range of other products and services that did not use CH data as an input.

Figure 3.5: Turnover generated by products and services that use CH data (as % of total turnover)



Source: ICF analysis of qualitative research with 15 intermediaries that use bulk data products.

[9] Total costs have been estimated based on costs of purchases, employment costs, taxes and capital expenditure.

Overall it was estimated that CH data was associated with products and services that generate revenues of £145m across the sample of intermediaries, representing around 16% of the total turnover of these businesses. However, the figures for the sub-samples show significant variation between the larger intermediaries (15% of their total turnover) and the mid-sized intermediaries (60% of their total turnover).

Applying the same metrics from the ABS provides combined estimates that the sample of intermediaries has incurred costs of £109m in providing these products and services.

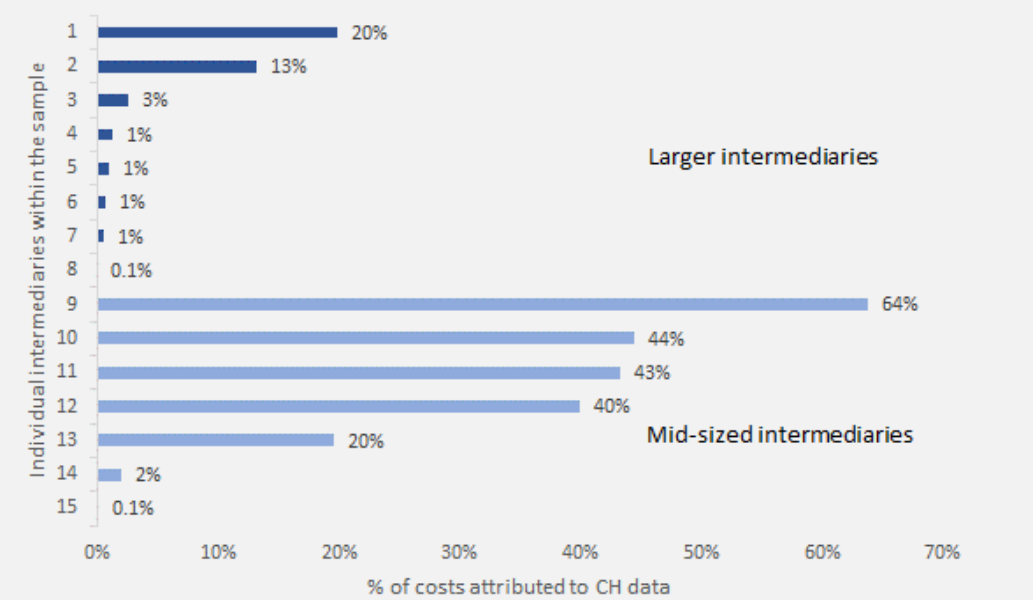
Table 3.1: Turnover and costs associated with intermediaries' total businesses and the products and services that use CH data

		Total business (£ value)			Products and services using CH data (£ value)		
		Larger intermed-iaries	Mid-sized intermed-iaries	Total sample	Larger intermed-iaries	Mid-sized intermed-iaries	Total sample
Turnover	£ value	£911m	£10m	**£921m**	£139m	£6m	**£145m**
Costs	£ value	£688m	£7m	**£695m**	£105m	£4m	**£109m**

Source: ICF analysis of qualitative research with 15 intermediaries that use bulk data products.

The sample of intermediaries was also asked if they could attribute the above costs and revenues to their specific use of CH data. Estimates again varied significantly between the different users, as shown below for costs (Figure 3.6) and revenues (Figure 3.7).

Figure 3.6: Attribution of costs to CH data (as % of costs of products that use CH data)

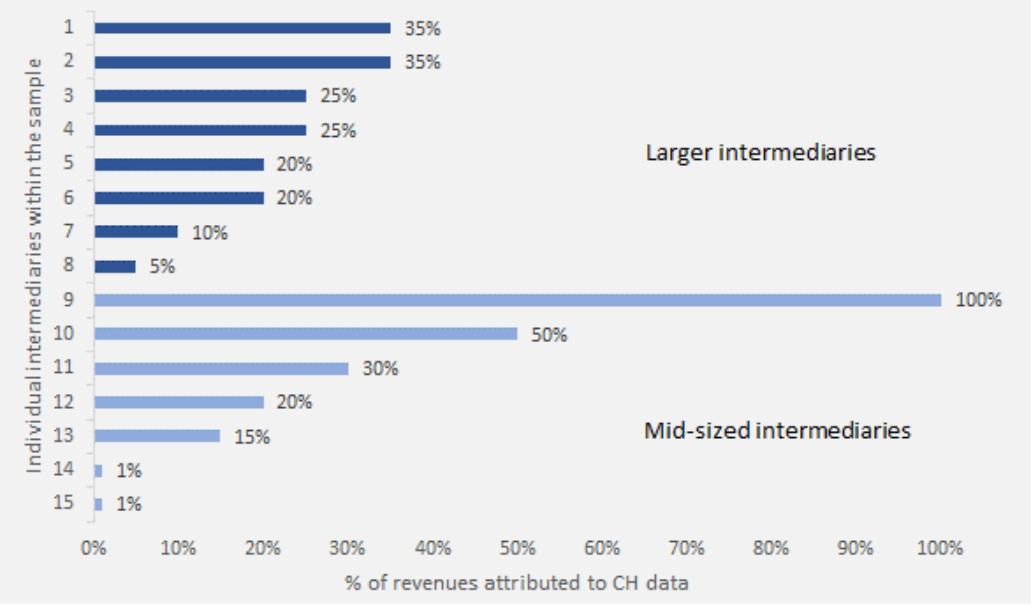


Source: ICF analysis of qualitative research with 15 intermediaries that use bulk data products.

Figure 3.6 shows considerable differences between the two sub-samples. The larger intermediaries reported that the costs attributed to CH data represented a relatively low proportion of the total costs of providing those products and services. Six of the eight larger intermediaries attributed less than 3% of these costs to CH data. This is unsurprising given the relatively low cost of purchasing CH data, much of which is available free of charge. However, this also includes the costs of cleaning, checking and processing the CH data and the higher costs attributed to CH data tended to be for the organisations that were manually entering large quantities of CH data that is only available as pdf images.

In contrast, the costs attributed to CH data were considerably higher for most of the mid-sized intermediaries. This was typically because CH data accounts for a larger share of the inputs for the products and services of these intermediaries, some of whom reported some relatively high costs for cleaning, checking and transforming the data before being input into their own datasets.

Figure 3.7: Attribution of revenues to CH data (as % of revenues of products that use CH data)



Source: ICF analysis of qualitative research with 15 intermediaries that use bulk data products.

Figure 3.7 shows the proportion of the revenues of those products and services that the intermediaries have attributed to CH data and highlights further key differences between the two sub-samples:

- Firstly, the larger intermediaries attributed a much larger proportion of revenues to CH data, relative to the costs (Figure 3.6). This suggests that the use of CH data enhances the productivity of these products and services by supporting greater revenues for a given unit of input.

- The larger intermediaries have also provided fairly consistent estimates of the revenues that they attribute to CH data, ranging from 5% to 35% of the total revenues of products using CH data.

- The mid-sized intermediaries provided much more varied estimates ranging from 1% to 100% of the total revenues of products that use CH data. The 100% figure is a unique case, involving a new business that is developing software to provide a user-friendly portal for exploring the CH datasets, which attributes 100% of the associated turnover to CH data.

Table 3.2 presents the combined cost and revenue estimates across the sample of intermediaries. It suggests that £23m of revenues can be attributed to CH data among the sample of intermediaries, which represents:

- 16% of the total revenues of products / services that use CH data as an input; and

- 2.5% of the total turnover of these businesses.

In contrast, the costs attributed to the use of CH data were considerably lower at £5.4m. This represents just 5% of the costs of producing products and services that use CH data and just 0.8% of the total costs of these businesses. This provides further evidence that the costs of using CH data are relatively low compared to the revenue that can be attributed to the data. As described above, while these estimates are only associated with the 15 intermediaries in the research sample, they are expected to include the most significant intermediaries, and are therefore likely to represent a large share of the total value of CH data across the total population of intermediaries.

Table 3.2: Attributing revenues and costs of intermediaries to the use of CH data

	Larger intermediaries		Mid-sized intermediaries		Total sample	
	£ value	% of products using CH data	£ value	% of products using CH data	£ value	% of products using CH data
Revenues	£21.5m	15%	£1.39m	22%	**£22.8m**	**16%**
Costs	£4.0m	4%	£1.36m	28%	**£5.4m**	**5%**

Source: ICF analysis of qualitative research with 15 intermediaries that use bulk data products.

Table 3.2 also shows some key differences between the larger and mid-sized intermediaries. It suggests that there is very little difference between the revenues (£1.39m) and costs (£1.36m) that are attributed to CH data across the mid-sized intermediaries. This was due to some relatively high costs of processing CH data among these businesses and some relatively low benefits of revenues reported by the sample. In contrast, the larger intermediaries attributed much larger revenues (£21.5m) to CH data, relative to the costs (£4m). This suggests that the benefits of CH data, in terms of

supporting revenues and productivity, were concentrated on the larger intermediaries within the sample.

Substitutes for CH data (the counterfactual)

The sample of intermediaries was also asked whether there were other products that they could use if CH data did not exist. Most of the sample did not feel there were any products that could substitute the data provided by CH. In the absence of CH data, most of the intermediaries suggested that they would either:

- remove certain data and functionality from their products and services, thereby reducing their quality and potentially reducing customer demand and revenues; or
- establish their own systems for collecting data directly from businesses. However, this was considered to be a very expensive option that would significantly increase costs, while resulting in lower quality data and less comprehensive coverage, which could also have a negative impact on revenues.

There was agreement across the sample that the absence of CH data would reduce the quality of their products and services, increase costs and/or reduce revenues.

Some of the intermediaries also described the loss of other benefits provided by CH data, in terms of credibility and legitimacy, given that CH data is widely regarded as the definitive source of company information in the UK. This was reported to enhance the reputation of the intermediaries' own products and services that use and provide access to CH data. These benefits would be lost in the absence of CH data, which could also have a negative impact on the revenues associated with these products and services.

However, most intermediaries in the sample were unable to provide quantified estimates of these impacts and there was also some variation in perceptions of the potential scale of impacts. For example, some intermediaries reported that these impacts could be significant and may result in changes to business models and, in two cases, potentially resulting in resources being redirected from the UK to international markets. In contrast, other intermediaries felt that impacts on revenues would be relatively minor as long as CH data was not available to any of their competitors and would only require some minor changes to their products and services.

Overall, however, it is likely that the lack of viable alternatives for much of the data provided by CH means that the true value of CH data for intermediaries is likely to be greater than the estimates of revenue attributed to CH data in the above analysis.

3.5 Suggested improvements to CH data

Most intermediaries in the sample reported that CH data has become more valuable over time following the introduction of new data products and services and more frequent updates to the data. It was also suggested that CH data has become more relevant following the financial crisis, which has made businesses and consumers more conscious of potential risks and more likely to want to be proactive in managing those risks.

However, several of the intermediaries also suggested that the introduction of free data had actually reduced the value of some of their own more basic products and services, since their customers could now access CH data free of charge for themselves. In these cases, the provision of free data and improved access to CH data has reduced costs for intermediaries but has also provided a new source of competition for some of their lower value-added products and services, which has had a negative overall impact on the benefits provided by CH data in relation to these particular products and services.

The fieldwork with intermediaries also identified several potential improvements to CH data. Enhanced powers for the Registrar are already proposed in many of these areas in BEIS' consultation on Corporate Transparency and Register Reform[10]. The potential improvements suggested by intermediaries related to the following themes:

- **Data transparency** – The interviews found that intermediaries have gained significant benefits from increased data transparency and the increased provision and breadth of CH data over time. They also have an appetite for even more data, and would like to see further increases in data provision and transparency, some of which could only be achieved through legislative changes to the powers of the Registrar (including some of those included in the Corporate Transparency and Register Reform consultation[11]). Suggestions included: addressing the requirements of the confirmation statement (that must be filed by all companies) so that companies are required to provide a full list of shareholders; increasing the consistency of reporting requirements across different types of entity, including small businesses (for example, while intermediaries understand the need for reduced requirements and exemptions for small businesses, they would also like to be able to access consistent accounts data across the whole of the UK business population); and continue to provide additional data and services (from CH, but also from other government departments such as HMRC).

- **Data quality** – The sample of intermediaries was generally satisfied with the overall quality of CH data but also suggested that there were opportunities for further improvements in data quality through additional data checking and quality assurance. The intermediaries reported that errors in the data are usually due to the

[10] BEIS (May 2019) Corporate transparency and register reform: Consultation on options to enhance the role of Companies House and increase the transparency of UK corporate entities. Available at: https://www.gov.uk/government/consultations/corporate-transparency-and-register-reform

[11] BEIS (May 2019) Corporate transparency and register reform: Consultation on options to enhance the role of Companies House and increase the transparency of UK corporate entities. Available at: https://www.gov.uk/government/consultations/corporate-transparency-and-register-reform

filing of inaccurate data, rather than issues with the processing of data, but some suggested that it would be helpful if CH could undertake more checks to identify errors and duplication and help improve data quality. One interviewee said that they would be willing to pay for CH to undertake greater quality control. It was also suggested that a more standardised approach to filing requirements would help to improve the quality of data.

- **Functionality of CH data products and services** – The interviews with intermediaries also identified suggestions around the functionality of CH datasets, including: a faster-streaming API service, providing access to raw data rather than pdf / image files (that several intermediaries were manually inputting into their systems); and undertaking more frequent reviews of data specifications and lookups within datasets as it was reported that there can sometimes be issues with these.

- **Communications between CH and intermediaries** – Some of the intermediaries also suggested there were opportunities to improve communications from CH regarding changes and issues with the data. Examples included: providing sufficient notice of changes to data files and formats (which can require intermediaries to amend and develop their own products and services to adapt to these changes), and informing users when there is likely to be a delay in the delivery of datasets (as many intermediaries operate a 'just-in-time' system where staff are lined up to clean and process CH data from the moment it is made available). Improving this type of communication would help these intermediaries to plan and allocate resources more effectively.

- **Consultation of key intermediaries in the development of new data products and services** – Two of the intermediaries suggested that they would welcome further opportunities to support CH in the development of new data products and services to ensure that they can meet the needs of intermediaries as effectively as possible. One user also suggested that they would welcome additional opportunities to engage and work with BEIS and CH and help to inform policy decisions regarding data transparency and the provision of public data.

4 Conclusions

This section summarises the key findings of the research and conclusions regarding the value of CH data to intermediaries that use CH data as an input to their own commercial products and services.

4.1 Estimating the value of CH data for intermediaries

Intermediaries are private sector businesses that that use CH data as an input to their own commercial products and services, thereby adding value to CH data and selling it on to end-users of the data. The research has found that this sub-group of users of CH data includes some very large businesses, such as credit reference agencies, but also includes much smaller organisations that are using CH data as an input to their own products or developing software and other tools to provide access to CH data.

It is difficult to estimate the total number of intermediaries in the UK, given the limited information with which to profile the overall population of users of CH data. The study therefore adopted two approaches to understanding the potential benefits of CH data for intermediaries:

- The WTP survey identified 85 businesses that reported including CH data in the products and services that they sell to their customers. These users represented 14% of the total number of survey responses, however, it is not possible to extrapolate to estimate the wider population of intermediaries due to the potential sample bias in the WTP survey[12].

- An alternative approach identifed the private sector businesses that access CH's 'bulk' data products. These products provide large quantities of CH data and users are likely to include the most significant intermediaries (i.e. those using very large quantities of CH data as inputs to their own products and services). CH identified 132 private sector businesses who are users of bulk data (i.e. intermediaries). There was no overlap between the the largest intermediary users of bulk data (that participated in the qualitative research) and the 85 responses to the WTP survey.

[12] The sampling approach for the direct user survey is described in Section 4.3 of Report 2. It was based around the 'total use' of CHS (frequency and duration) and did not attempt to control for user type (e.g. direct users, intermediaries, providers of public goods). Hence it is difficult to conclude that the observed incidence of intermediaries in the survey (14% of the sample) is a reliable result, and therefore we would not recommend aggregating results on this basis.

The analysis of the research findings suggests that there are three broad, and potentially overlapping, categories of intermediary. These are:

- A group identified through the WTP survey of typically 'smaller' intermediaries (i.e. approximately half have an annual turnover of less than £250,000 and/or employ between 0-4 employees) whose willingness to pay is similar to the direct users of CH data. The responses to the WTP survey suggested that the average annual benefit to these beneficiaries is approximately £2,000 per user per year, and an average benefit of approximately £8 per individual 'use' of CH data. These figures are consistent with the overall estimates for all direct users, presented in Report 2.

- A group of 'mid-sized' intermediaries, which in our research sample, have an average turnover of just over £1m each and where, on average, 60% of their turnover depends on products and services that use CH data. In the research sample, the seven 'mid-sized' intermediaries attributed £1.4m of their revenues to the use of CH data.

- A group of 'larger' intermediaries which, in our research sample, have an average turnover of over £100m each. This group includes the main credit reference agencies. These are large multi-product businesses, for which sales generated by products and services that use CH data are on average 15% of their total turnover. In our sample, the eight 'larger' intermediaries attributed £21.5m of their revenues to the use of CH data (i.e. representing 94% of the total 'mid-sized' and 'larger' intermediary revenues attributed to the use of CH data).

Collectively the research sample of 15 'mid-sized' and 'larger' intermediaries directly attributed £23m of their revenues to the use of CH data and £5m of their costs to accessing and processing CH data, (i.e. contributing net income of around £18m[13]).

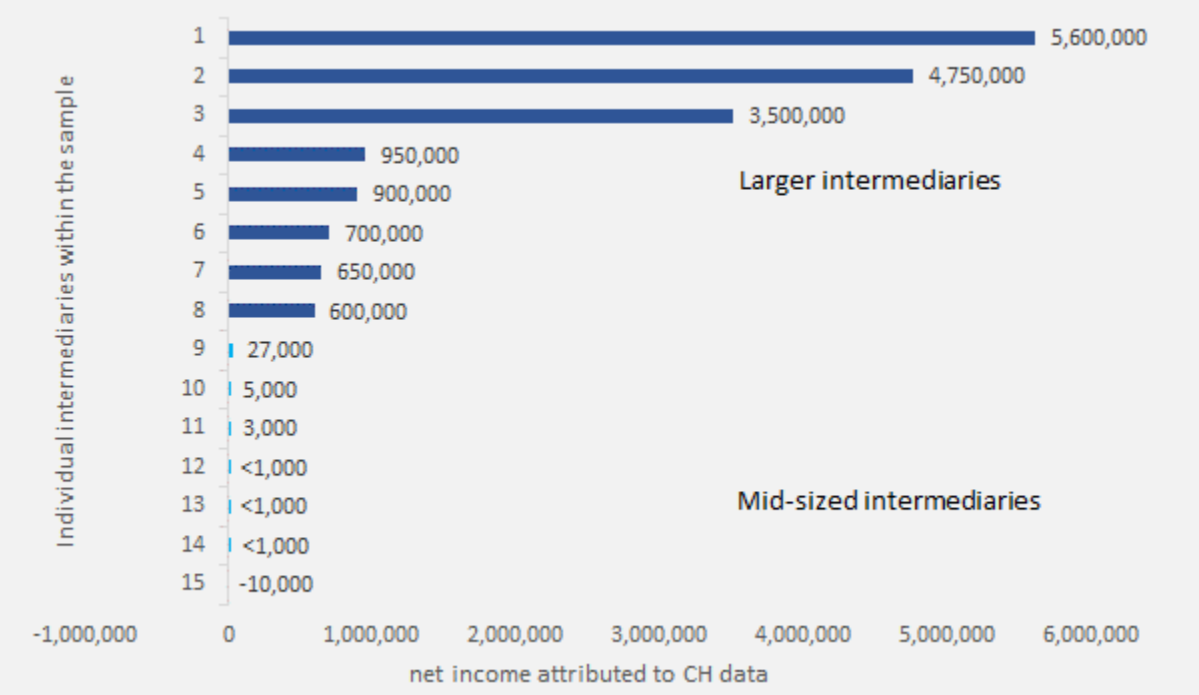
However, the variation in estimates between individual businesses was very large, This is illustrated in the figure below which shows the distribution of net income from CH data for the sample. It shows that net income estimates ranged from -£10,000 (in the case of a new business that is currently developing software to explore the CH datasets, and expects future revenues to increase significantly) to £5.6m in the case of one of the largest intermediaries. There was also a significant difference between the two sub-samples:

- The 'mid-sized' intermediaries attributed an average net income of less than £5,000 to the use of CH data, with individual estimates ranging from -£10,000 to £27,000 per year.

[13] However, the net income figure is unlikely to provide an accurate estimate of the change in producer surplus (as described in Report 1), as it is unlikely that all intermediaries in the sample have consistently and appropriately attributed their revenues and costs to the use of CH data.

- The 'larger' intermediaries' attributed an average net income of approximately £2.2m to the use of CH data, with individual estimates ranging from £0.6m to £5.6m.

Figure 4.1: Net income attributed to CH data



Source: ICF analysis of qualitative research with 15 intermediaries that use bulk data products.

The above estimates provide average figures for net income that are significantly larger than the £2,000 per user that was identified in the WTP survey for 'smaller' intermediaries. This also suggests that the benefits for intermediaries that use other 'non-bulk' products and services are likely to be much lower in value than the benefits for the intermediaries that use bulk data products.

The qualitative research also identified additional benefits of CH data in terms of:

- being able to use company numbers in the CH data as a means of linking CH data with other sources of data (and in many cases allowing intermediaries to develop systems to automate this matching process); and

- the credibility and legitimacy provided by CH data, which is widely regarded as the definitive source of company information in the UK and enhances customer perceptions and the reputation of intermediaries' own products and services that use CH data.

The intermediaries also reported a lack of viable substitutes for CH data. In the absence of CH data, most of the intermediaries suggested that they would either have to remove certain data and functionality from their products and services, or develop their own systems for collecting data directly from businesses. While the intermediaries were unable

to quantify or monetise these impacts, they reported that the absence of CH data would be expected to reduce the quality of their products and services, increase their costs and/or reduce their revenues.

4.2 Changes in benefit values over time

More than half of the intermediaries that access CH bulk data products have only been accessing these products since they became available free of change. This suggests that most intermediaries are relatively new users of these products.

However, this was in contrast to the qualitative research sample, which included the largest intermediaries, who tended to be longer-term users of CH data. Most of these intermediaries suggested that their usage of CH data had remained constant over time and the introduction of free data had only had a minimal impact on their levels of usage. For most of these customers, the introduction of free data had provided cost savings, but was unlikely to have affected their levels of use. However, the large numbers of new users among the wider population of bulk data users suggests that the introduction of free data may have had a greater impact for these other 'mid-sized' intermediaries.

The qualitative research also identified some negative impacts for intermediaries from the introduction of free data in terms of providing a new source of competition for some of their lower value-added products and services, which customers can access for free themselves, directly from CH. This has reduced the value of some of these more basic products and services for intermediaries.

Overall, the qualitative research found that mid sized and larger intermediaries were accessing more CH data over time, but this was generally due to the additional of new datasets, and the expansion and increased frequency of existing datasets, rather than the introduction of free data. The introduction of PSC data has been a particularly significant addition, enabling some intermediaries to improve existing products and develop new products and services.

4.3 CH data that generates the greatest user value

The sample of intermediaries reported using a broad range of different CH products and services. The most commonly used and most valued pieces of CH data among the sample of key intermediaries (who use bulk data products) were:

- the CHS-API, for the ability to gain direct access to CH data and systems;
- bulk data, due to the breadth of data available and the ability to run searches; and
- PSC data, for providing access to company owners that tends to be difficult to access from other sources.

In contrast, the intermediaries identified in the WTP survey were more likely to use the CHS, WebCHeck and Companies House Direct services.

It was also suggested that CH data could be further improved to add even greater value by:

- improving data transparency – for example, by ensuring companies are required to provide a full list of shareholders, increasing the consistency of reporting requirements across all types of entity, and continuing to increase the overall provision of data and services from CH and other government departments (such as HMRC).

- making further improvements to data quality – for example, through additional data checking and quality assurance to identify errors and duplication.

- improving the functionality of CH data – for example, by providing a faster-streaming API service, providing access to raw data (rather than pdf / image files), and undertaking more frequent reviews of data specifications and lookups within datasets.

- increasing communications – for example, by providing sufficient notice of changes to datasets and formats, and informing users of likely delays in the delivery of datasets, to help intermediaries plan and allocate resources more effectively.

- increasing consultation with intermediaries – for example, to support CH in the development of new data products and services to ensure they meet the needs of users, and to help inform policy decisions regarding data transparency and the provision of public data.

ANNEXES

Annex 1 Interview topic guide for mid-size and largest 'intermediaries'

General information

ASK ALL:

- Please can you start by providing some brief information about your business, in terms of its size and the type of work that it does. *Probe:*
    - *Sector and main activity of business*
    - *Size of UK business (Number of employees / turnover)*
    - *Approximately how long the business has been trading for*
- Please can you also tell me a little bit about your role and your experience of accessing and using data from CH. *Probe:*
    - *Role of interviewee*
    - *Extent of experience of using CH data (within current and/or previous roles / companies), including length of experience of accessing CH data*

Use of CH data

ASK ALL:

- How long has your business been accessing data from CH or when did you start accessing data from CH?
- Why does your business access CH data? What is the CH data used for?
- Do you use CH data to develop your own products and services? If so:
    - What are the products and services that make use of CH data? Please can you describe.
    - Who are the customers of these products and services? *Probe the type of business, organisation or individual, numbers of customers*
    - How much revenue is generated by these products and services?

In order to route the interviewee through relevant questions, please ask the following questions to understand the type of data being accessed:

- Do you pay fees for any of the data that you access from CH?

- Do you use any of the free data products and services provided by CH?
- Do you access data from CH on People with Significant Control (PSC)?

ASK THOSE PAYING FEES FOR CH DATA:

- Which of the following data products and services do you pay a fee to access from CH? *Probe for all products and services in the table below. If the interviewee is not aware of the different products/services, ask if they can describe the CH data that they pay for.*

'Paid-for' data products / services	Notes for interviewer		
	Description	Fees	Registration required
DVD directory	Monthly DVD of CH data with built-in search facility (2 versions: 1 allows export of data, the other does not)	Yes (monthly / annual fees, plus additional cost for export version)	Yes
XML gateway search service	A PC to PC service that allows users to search CH data from their own software / office	Yes, monthly subscription (plus fees for additional docs / requests)	Yes
Contact centres	Tel / email requests for data submitted to CH contact centres	Yes	Yes, to receive info
Information centres	Sites at Cardiff, Belfast, Edinburgh and London that can be used by the public to access data	Not for basic data. Yes for more detailed data / services	Not unless required to receive info
(Purchasing documents or images from) WebCHeck	Web-based search facility	No (except to purchase docs or images)	Not for basic data. Yes for more detailed info
Information not on the public register	Specified public authorities (SPAs) and credit reference agencies (CRAs) can pay to access additional information not on the public register	Yes (applications = £54; individual requests = £5)	Yes

- Does your business pay for any other data or services from CH? If so, please can you describe these other products or services.
    - *Note for interviewer: This could include 'bespoke products' (such as bespoke data on liquidations, insolvencies, forms filed, etc.) or requests for management information that is not available from the other data products.*
    - *Probe: name / type of product or service, the data provided, the frequency of use and the cost.*

For each 'paid for' product that they use, please ask:

- How frequently does your business access this product?

- How much do you pay to access this product?
    - *Ensure you record detail of costs per month / year (total costs and per product / service / type of output) and the type of costs (e.g. regular subscription costs; costs for particular products / services; and ad hoc / variable costs for particular outputs)*
- What is this product used for? *Probe: does it help them to develop their own products / services, if so, which ones and how?*

ASK THOSE ACCESSING FREE DATA:

- Which of the following free data products and services do you access from CH? *Probe for all products and services in the table below. If the interviewee is not aware of the different products/services, ask if they can describe the free data that they access.*

Free data products / services	Notes for interviewer		
	Description	Fees	Registration required
Companies House Service (CHS)	Web-based search facility of real-time data	No	No, unless receiving email alerts
Companies House Service (CHS) – API service	CHS data (as above) but in a format that can be searched using other software	No	Yes, to receive data via email
WebCHeck	Web-based search facility	No (except to purchase docs or images)	Not for basic data. Yes for more detailed info
Company data product	Monthly snapshot for bulk data download (basic details)	No	No
Accounts data product	Downloadable zip-file of filed accounts (daily/monthly data)	No	No
People with Significant Control (PSC) data product	Downloadable daily snapshot of all listed PSCs	No	No
Uniform resource identifiers (URI)	Service providing URIs (unique URLs) for each company listed with CH	No	No
Mobile app	Search facility for Android and iOS devices	No	No
Information centres	Sites at Cardiff, Belfast, Edinburgh and London that can be used by the public to access data	Not for basic data. Yes for more detailed data / services	Not unless required to receive info

- Does your business access any other free data or services from CH? If so, please can you describe these other products or services.
    - *Note for interviewer: This could include requests for bespoke data and/or management information that is not available from the other data products.*
    - *Probe: name / type of product or service, the data provided, the frequency of use*
- How frequently does your business access each of those free data products? *Probe for all 'free' products that they are currently using*
- How does your business use the free data provided by CH? *Probe: are these used to help them to develop their own products / services, if so, which ones and how?*
- Has the increased availability of free data affected your business in any way? If so, how and why? *Probe: increased use of CH data, reduced costs / increased productivity / efficiency*

ASK THOSE ACCESSING PSC DATA:

- Please can you tell me how your business accesses data from the PSC register.
- How frequently does your business access PSC data from CH?
- What is the PSC data used for? Is the PSC data used to develop any of the products and services that you have described already? If so, which?
- Has the PSC data been used to add functionality to any of these existing services or to develop any bespoke products or services? If so, please describe.
- Has the provision of PSC data affected your business in any way? If so, how and why? *Probe: increased use of CH data, reduced costs / increased productivity / efficiency*

Key trends and future opportunities

ASK ALL:

- Has there been any change over time in the CH data products that your business accesses or the frequency of access?
    - If so, how, when and why has it changed? *Probe: extent to which this is linked to when products became free to access or when new products or services were introduced (e.g. PSC data)*
    - Do you think that the data you access from CH has become more or less valuable over time? If so, why? *Probe: the extent to which changes are due*

to the type/range of products/services available, the introduction of free/PSC data

- Are there particular CH data products that are most useful and/or add most value for your business? If so:
    - Which CH products are most useful and why?
- Are there any ways in which the provision of CH data could be improved to make it easier for your business to use?

Attribution of costs and revenues to CH data

ASK ALL:

- Do you have to process, clean or do anything else with the data you access from CH before you are able to use it? If so, please can you:
    - Describe what is involved in these activities and the reasons for doing so.
    - State which CH data products are processed or cleaned in this way.
    - Estimate the approximate time spent and any costs incurred in these activities.
- I would now like to ask you about the fixed costs you face in providing products and services that use CH data. Thinking of all of the costs involved in producing the products and services that use CH data… approximately what proportion of these costs would you say are fixed costs, as opposed to variable costs?
    - *Note to interviewer: fixed costs can be defined as: costs that do not increase or decrease with the number of products and services produced and sold, such as:*
        - *costs of staff time and purchases of data and other materials that are incurred regardless of the level of production or sales; and*
        - *other fixed costs including rent, insurance, taxes, utilities, interest payments, asset depreciation.*
    - *(If unable to provide a figure), probe for an approximate figure using the following categories: 0%-20%; 21%-40%; 41%-60%; 61%-80%; or 81%-100%*
- And approximately what proportion of these costs would you say is associated with the purchase and processing of data from CH?
    - *(If unable to provide a figure), probe for an approximate figure using the following categories: 0%-20%; 21%-40%; 41%-60%; 61%-80%; or 81%-100%*

ASK ALL:

- More generally, how important would you say CH data is to the development and provision of these products and services? *Probe why*

- Based on your responses to those questions, approximately what proportion of the revenues generated by these products would you attribute to the data provided by CH?

    - *(If unable to provide a figure), probe for an approximate figure using the following categories: 0%-20%; 21%-40%; 41%-60%; 61%-80%; or 81%-100%*

    - Why do you say that? Please explain

Current use of complementary data sources

ASK ALL:

- Do you use any other data sources in combination with CH data in developing your own products and services? If so:

    - What other data sources are used to develop these products and services?

    - Why do you use these different data sources? What information does each source provide?

    - What are the key differences between the data provided by CH and other sources in terms of the type of information provided, the way it is presented/accessed, the quality of the data/product/service?

        - Do the other sources provide information that is not available from CH? If so, what? *Probe for extent of overlap between CH data and other sources and extent to which they provide additional data*

    - How easy or difficult is the processing of CH data to enable it to be combined with these other sources? Why do you say that?

Availability of substitutes for CH data

ASK ALL:

- To what extent could your products and services be provided in the absence of CH data? *Probe to what extent they would have been able to develop these products and services (and establish these business functions) in the absence of CH data.*

- If CH data did not exist, are there other sources of data that you could use?

    If so:

    - What are these alternative sources?

- How does the data available from these alternative sources differ from the data provided by CH? Is it of a similar quality to CH data? If not, please describe any differences.

- Please can you estimate the time and costs associated with accessing and processing data from these alternative sources? How might the time and costs be expected to differ from those associated with accessing and processing CH data? *Probe if it is possible to quantify these estimates*

- Do you think there would be any changes in the prices you would charge or the revenues received from these products / services? If so, please describe.

- Why do you choose to use data from CH? Are there benefits provided by CH data, compared to these alternative sources? If so, what? *Probe: cost savings, greater efficiency, better quality, greater breadth of information*

If not,

- Are there any other ways of accessing the data and type of information that you are currently accessing from CH?

  If yes, probe:

  - *what this might be expected to involve;*
  - *how long it would take to access/produce the required information, and the estimated costs of doing so; and*
  - *what would be the impacts for your business (and any end users) of accessing data in this way, rather than using data from CH (e.g. higher costs, lower efficiency, lower quality products and services, or things that they would not be able to provide).*

  If no, probe:

  - *what their business would do in the absence of CH data; and*
  - *the resulting impacts for their business and/or their end users (e.g. would they make changes to their products/services, are there any that they would no longer be able to provide; what would be the impact of not providing this information or these services?)*

Closing remarks

ASK ALL:

Are there any other points you would like to make about CH data that have not already been discussed?

Annex 2 Online survey questionnaire for 'intermediaries'

About your organisation

1. What sector does your business operate within?

- Administrative and support services
- Financial services
- Information and communication services
- Legal and accounting services
- Public sector
- Other (please specify)

2. How many people are employed by your business?

- 0-9 people
- 10-49 people
- 50-249 people
- 250+ people

3. What is the annual turnover of your business?

Companies House data

4. Which of the following data products and services do you access from Companies House?

- Companies House Service (CHS)
- Companies House Service (CHS) – API service
- WebCHeck
- Company data product

- Accounts data product
- People with Significant Control (PSC) data product
- Uniform resource identifiers (URI)
- Mobile app
- Information centres
- DVD directory
- XML gateway search service
- Contact centres
- Information not on the public register
- Other bulk data
- Bespoke data
- Other (please specify)

Costs of using Companies House data

5. How much do you pay to access Companies House data?

6. Do you have to process, clean or do anything else with the data you access from Companies House before you are able to use it?

- Yes
- No

6a. *(If yes)* Please estimate the approximate time spent and any costs incurred in these activities.

Products and services using Companies House data

7. How much revenue does your organisation generate by selling products and services that make use of Companies House data? Please specify the value in pounds and whether it is per week, month or year.

8. Approximately what proportion of those revenues would you attribute to the data provided by Companies House?

If you're unable to provide a figure, please select from one of the ranges below: *This question is required.

- Less than 1%
- 1%-3%
- 4%-9%
- 10%-19%
- 20%-49%
- 50% and above

9. Approximately what proportion of the costs of these products and services would you say is associated with the purchase and processing of data from Companies House?

If you're unable to provide a figure, please select from one of the ranges below: *This question is required.

- Less than 1%
- 1%-3%
- 4%-9%
- 10%-19%
- 20%-49%
- 50% and above

Availability of substitutes for Companies House data

10. If Companies House data did not exist, are there other sources of data that you could use?

- Yes
- No

11. *(If yes)* How are the data available from these alternative sources different from the data provided by Companies House (in terms of quality)?

The data available from these other sources is:

- Higher quality

- Same quality
- Lower quality

12. *(If yes)* How are the data available from these alternative sources different from the data provided by Companies House (in terms of price)?

The data available from these other sources is:

- More expensive
- A similar price
- Less expensive

13. *(If yes)* Do you think the time and costs of accessing and processing data from alternative sources would be…?

- higher than those associated with Companies House data
- the same as those associated with Companies House data
- lower than those associated with Companies House data

14. *(If higher or lower)* How much higher or lower do you think the time and costs would be, if you were using these alternative sources?

If you're unable to provide a figure, please choose from the following below categories:

- More than 100% higher
- 81%-100% higher
- 61%-80% higher
- 41%-60% higher
- 21%-40% higher
- 0%-20% higher
- No Change
- 0%-20% lower
- 21%-40% lower
- 41%-60% lower

- 61%-80% lower
- 81%-100% lower

Closing remarks

15. Do you have any other comments you would like to make regarding the benefits of Companies House data for their organisation?

Would you be happy to be contacted by ICF for a follow-up telephone interview to gain a more detailed understanding of the benefits of Companies House data for your organisation? Please note that we may not contact you, and if we do, we will first make contact by email to confirm that you still wish to participate and to provide further details on the purpose of the interview.

- Yes
- No

(If yes) Please provide the following information:

- Name:
- Company name:
- Telephone number
- Email

This publication is available from: www.gov.uk/government/publications/companies-house-data-valuing-the-user-benefits

If you need a version of this document in a more accessible format, please email enquiries@beis.gov.uk. Please tell us what format you need. It will help us if you say what assistive technology you use.